Orthogonal Frequency Division Multiplexing for Wireless Communications

A Thesis
Presented to
The Academic Faculty

by

Hua Zhang

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Electrical and Computer Engineering

School of Electrical and Computer Engineering
Georgia Institute of Technology
November 11, 2004

Orthogonal Frequency Division Multiplexing for Wireless Communications

Approved by:

Professor Gordon L. Stüber,
Committee Chair, Electrical and Computer Engineering

Professor Ye (Geoffrey) Li, Advisor
Electrical and Computer Engineering

Professor Guotong Zhou,
Electrical and Computer Engineering

Professor Gregory D. Durgin,
Electrical and Computer Engineering

Professor Xinxin Yu,
School of Mathematics

Date Approved: November 16, 2004

To my parents,

Zhang Zhongkang and Gong Huiju

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere thanks to my advisor, Dr. Geoffrey (Ye) Li, for his support, encouragement, guidance, and trust throughout my Ph.D study. He teaches me not only the way to do research but the wisdom of living. He is always available to give me timely and indispensable advice. I can never forget the days and nights he spent on my papers. Three years is short compared with one's life, but it is enough to change one's whole life. What I learned within the three years working with Dr. Li established a basis that could lead me to the future success. All I could say is that I can never ask for any more from my advisor.

Next, I would like to thank Dr. Gordon L. Stüber, Dr. Guotong Zhou, and Dr. Gregory D. Durgin for serving on my proposal committee. Their insightful comments on my proposal helped me to finish the research towards this thesis. Furthermore, I would like to thank the former and current members of Information Transmission and Processing Laboratory (ITP), Guocong Song, Jianxuan Du, Jet Zhu, Jingnong Yang, Taewon Hwang, Uzoma Anaso Onunkwo, and Ghurumuruhan Ganesan for their help, encouragement and the fruitful discussions. They create a great research environment. Without their help, the thesis is impossible. Their friendship gives me courage and warm to overcome the loneness during my study.

Finally, I would like to thank my parents for their support and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Orthogonal frequency division multiplexing (OFDM) is a promising technique for high-data-rate wireless communications because it can combat inter-symbol interference (ISI) caused by the dispersive fading of wireless channels. The proposed research focuses on techniques that improve the performance of OFDM based wireless communications and its commercial and military applications. In particular, we address the following aspects of OFDM: inter-carrier interference (ICI) suppression, co-channel interference suppression for clustered OFDM, clustered OFDM based anti-jamming modulation, channel estimation for MIMO-OFDM, and precoding for MIMO-OFDM with channel feedback.

For inter-carrier interference suppression, a frequency domain partial response coding (PRC) scheme is proposed to mitigate ICI. We derive the near-optimal weights for PRC that are independent of the channel power spectrum. The error floor resulting from ICI can be reduced significantly using a two-tap or a three-tap PRC.

Clustered OFDM is a new technique that has many advantages over traditional OFDM. In clustered OFDM systems, adaptive antenna arrays can be used for interference suppression. To calculate weights for interference suppression, we propose a polynomial-based parameter estimator to mitigate the severe leakage of the discrete Fourier transform (DFT)-based estimator due to the small size of each cluster. An approximately optimal window size for the polynomial-based estimator is obtained and an adaptive algorithm is developed to obtain the optimal window size. With the adaptive algorithm, the polynomial-based estimator has no leakage and does not require channel statistics.

Clustered OFDM can also be applied to military communications for high-data-rate transmission. We propose a clustered OFDM based spread spectrum modulation to provide better anti-jamming capability. The approximate and asymptotic expressions are derived for performance analysis. For coded systems with hard- and soft-decision decoding, the performance bounds are used to evaluate the anti-jamming performance. We have also developed a simple jamming state estimator for soft-decision decoding.

Employing multiple transmit and receive antennas in OFDM systems (MIMO-OFDM) can increase the spectral efficiency and link reliability. However, channel estimation is a challenging task for MIMO-OFDM systems since more parameters need to be estimated than in single transmit antenna systems. We develop an minimum mean-square-error (MMSE) channel estimator that takes advantage of the spatial and frequency correlations in MIMO-OFDM systems to minimize the estimation error. We derive conditions for the optimal training sequences and investigate the training sequence designs for several channel conditions. Two optimal training sequence designs for arbitrary spatial correlations are developed. The training sequence designs for some special correlations are also discussed.

For a MIMO system, the diversity and array gains can be obtained through optimal linear precoding if the exact knowledge of channel state information is available at the transmitter. In practical implementations, perfect channel state information is difficult to be obtained. We propose to use a linear precoding scheme that requires limited feedback. We investigate the issues of codebook construction and selection criteria. For MIMO-OFDM systems, we propose a subspace tracking based approach that can exploit the frequency correlations between the OFDM subchannels to reduce the feedback rate. The proposed approach does not require recalculation of the precoding matrix and is robust to multiple data stream transmissions.

CHAPTER I

INTRODUCTION

The growth of mobile communications and wireless Internet access has produced a strong demand for advanced wireless techniques. The challenges for wireless communication designs come from the detrimental characteristics of wireless environments, such as multipath fading, Doppler effect, co-channel interference, and intentional jamming in military communications. The objective of our research is to provide new approaches to solve the problems mentioned above by means of *orthogonal frequency division multiplexing* (OFDM).

Multipath fading of wireless channels leads to *inter-symbol interference* (ISI), which limits the transmission rate of single-carrier systems. In conventional single-carrier communication systems, the ISI is usually dealt with by a time domain channel equalizer [1]. When the data rate increases, the symbol duration reduces and the equalizer becomes very complex. OFDM is an elegant solution to the severe ISI problem [2], [3].

OFDM is a special form of multicarrier modulation [4], which was originally used in high frequency military radio. An efficient way to implement OFDM by means of a *Discrete-time Fourier Transform* (DFT) was found by Weinstein in 1971 [2]. The computational complexity could be further reduced by a *Fast Fourier Transform* (FFT). However, OFDM was not popular at that time because the implementation of large-size FFTs was still too expensive. Recent advances in VLSI technologies have enabled cheap and fast implementation of FFTs and IFFTs. In the 1980s, Cimini first investigated the use of OFDM for mobile communications [3]. Since then, OFDM has become popular. In the 1990s, OFDM was adopted in the standards of *digital audio*

broadcasting (DAB), *digital video broadcasting* (DVB), *asymmetric digital subscriber line* (ADSL), and IEEE802.11a. OFDM is also considered in the new fixed broadband wireless access system specifications.

In OFDM systems, the entire channel is divided into N narrow subchannels and the high-rate data are transmitted in parallel through the subchannels at the same time. Therefore, the symbol duration is N times longer than that of single-carrier systems and the ISI is reduced by N times. Through adding a *cyclic prefix* (CP) ahead of each OFDM symbol, the ISI can be totally suppressed as long as the length of CP T_g is longer than the maximum channel delay τ_{max}. Usually the length of the cyclic prefix is much smaller than the symbol duration, therefore, the spectrum efficiency decrease is negligible. To preserve the orthogonality, the subchannel spacing satisfies $\Delta f = 1/T_s$, where T_s is the OFDM symbol duration.

OFDM modulation and demodulation can be efficiently implemented by an IFFT and FFT. Although OFDM successfully prevents the ISI, it does not suppress channel fading. By using coding and interleaving across the frequency and time domain, the transmitted data can be effectively protected. Further improvement can be achieved through other advanced techniques, such as power allocation and adaptive modulation. Since the different subchannels experience different fading in the frequency selective channels, the optimal power allocation that maximizes the total capacity is water pouring [5], i.e., allocating more power to subchannels with high gains. Adaptive modulation is a simple way to combat the deep fading in some subchannels. For adaptive modulation, the constellation size of modulation for each subchannel is adjusted according to the subchannel quality such that a low bit error rate is preserved. OFDM can also be used together with multiple access schemes, where the subchannels, power, and data rate are dynamically allocated to provide a high degree of flexibility in supportable bit rates and Quality-of-Service (QoS) [22], [40].

OFDM has many good properties that make it an attractive modulation for high-data-rate transmission. However, it has also some inherent disadvantages. One of its disadvantages is the large *peak-to-average power ratio* (PAPR). Theoretically, the difference of the PAPR between an OFDM system and a single carrier system is proportional to the number of subchannels, though the theoretical value rarely happens. Large PAPR reduces the efficiency of the power amplifier, and results in nonlinear distortion of the transmitted signal. Several techniques have been proposed to reduce the PAPR, such as clipping, coding, peak windowing, phase shifting, and so on.

Another disadvantage is that OFDM is sensitive to Doppler frequency and carrier offset, because the bandwidth of each subchannel is very narrow. Because the subchannels are closely spaced, the orthogonality among subchannels is destroyed by time variation over one OFDM symbol or carrier frequency offset [16]-[19]. This causes *inter-carrier interference* (ICI). If not compensated for, the ICI will result in an error floor, which increases with Doppler frequency and symbol duration. We propose a frequency domain *partial response coding* (PRC) to reduce the effect of the ICI. Based on the general expression of the ICI power for OFDM with PRC, we derive the near-optimum weights for PRC that minimize the ICI power. From the numerical and simulation results, PRC with optimal weights for OFDM can reduce the ICI effectively.

Recently, a novel technique, referred to as clustered OFDM [20]-[22], was proposed to improve the performance of classical OFDM systems. In a clustered OFDM system, the wideband OFDM signal is organized into clusters of subchannels in frequency domain. Each user can access several clusters located at different frequencies. If channel coding is used over the clusters, frequency diversity gain can be obtained. Clustered OFDM also provides a flexible multiple access scheme for multiuser communications. Through a simple allocation algorithm, the whole system performance can be improved [40].

In multiuser communication systems, such as cellular systems, the performance is limited by co-channel interference. Adaptive antenna arrays have been proven to be an effective technique to mitigate fading effect and suppress co-channel interference, thereby increase the link reliability and coverage of wireless communications [41]-[43]. Among many approaches for interference suppression, *minimum mean-square-error diversity combing* (MMSE-DC) is the most effective one. To calculate weights for the MMSE-DC, the receiver needs the channel state information of the desired signals and statistics of co-channel interference, which are obtained through estimation in practice [45]. We investigate adaptive antenna arrays for clustered OFDM to suppress co-channel interference. Due to the small size of each cluster for clustered OFDM, the DFT-based estimator [45] has large leakage and results in severe performance degradation. Therefore, a polynomial-based parameter estimator is proposed to combat the severe leakage of the DFT-based estimator. We study the impact of polynomial order and window size on the estimation error. An approximately optimal window size for the polynomial-based estimator is derived and an adaptive algorithm for the optimal window size is developed. With the adaptive algorithm, the polynomial-based estimator has no leakage and does not require channel statistics.

Clustered OFDM can be also applied in military communications to design a *low probability interception* (LPI) and anti-jamming modulation. We have investigated clustered OFDM for military communications [62]. For military applications, the system has to be designed to protect against intentional interference, jamming. *Spread spectrum* (SS) is known to be an effective anti-jamming technique [58]. In a spread spectrum system, the transmitted signal is spread over a wide frequency band, much wider than the minimum bandwidth required to transmit data. The transmitter spreads the signal over a large bandwidth through a pseudo-random code known to the intended receiver. The receiver can despread the received signal using the same pseudo-random code. Since the jammer does not know the pseudo-random code,

the jamming signals only slightly increase the noise floor at the receiver. However, many military applications need high-data-rate transmission over wireless channels, which requires OFDM to deal with delay spread. The combination of OFDM and spread spectrum, so-called *multicarrier spread spectrum* (MC-SS) [59], has better anti-jamming and ISI suppression performance than single-carrier SS. It was shown in [60] and [61] that MC-SS outperforms single-carrier DS-SS. We combine clustered OFDM with spread spectrum techniques to design an anti-jamming modulation for military communications. We analyze the anti-jamming performance of clustered OFDM spread spectrum systems for dispersive channels. Since there is no close-form expression for the multi-channel gain output distribution with arbitrary correlation matrices, we use an exponential correlation matrix to approximate the practical channel correlation matrix. The approximate performance is very close to the exact one. For coded systems with hard- and soft-decision decoding, we investigate their performance bounds. Our numerical results show that the proposed scheme provides great ability of anti-jamming and combating the dispersive fading of wireless channels.

Multiple transmit and receive antennas can be used to form *multiple input and multiple output* (MIMO) channels and increase channel capacity. For a narrowband communication system with M_T transmit antennas and M_R receive antennas in rich scattering propagation environments, it is shown [47]-[48] that the information-theoretical capacity increases linearly with the minimum number of transmit and receive antennas, $\min(M_T,\ M_R)$. Various schemes have been proposed to exploit the advantages of MIMO channels, such as space-time coding [49] and BLAST [53]. Employing multiple antennas in OFDM systems (MIMO-OFDM) can reduce the equalizer complexity for broadband communication systems. However, most of MIMO transmission and signal detection schemes require channel state information. In MIMO systems, channel estimation is more challenging than in single antenna systems

since more parameters have to be estimated [55]. Most analysis and channel estimations assume that MIMO channels are *independent and identically distributed* (i.i.d.) Rayleigh fading. In indoor environments, however, MIMO channels are correlated [56], [57] and with Ricean fading. Exploiting this characteristics, channel estimation can achieve better performance. We develop an MMSE channel estimator for MIMO-OFDM systems that can makes full use of the spatial and frequency correlations and design optimum training sequences that minimize the channel estimation error. In general, the optimal training sequences for different transmit antennas are orthogonal and with equal power. In certain special cases, the power of training sequences can be further optimized.

For MIMO systems, both diversity and multiplexing gain can be obtained simultaneously [54]. Close-loop method can provide both diversity and multiplexing gain with low complexity. However, close-loop method requires accurate channel state information. In practical implementations, it is difficult for the transmitter to obtain perfect channel state information. We investigate a precoded MIMO system with limited feedback, whereby a precoding codebook is constructed to quantize the MIMO channel subspace. The receiver conveys only the indices of the best precoding matrix to the transmitter. We discuss the precoding matrix selection problem and propose simplified sub-optimum algorithms. For MIMO-OFDM systems, we investigate the clustering, interpolation, and subspace tracking approaches. The proposed subspace tracking approach can reduce the feedback rate and is robust to multiple data stream transmissions.

The rest of the thesis is organized as follows. In Chapter 2, we study ICI suppression for OFDM. We give details on the frequency domain partial response coding. In Chapter 3, two issues for clustered OFDM are addressed: co-channel interference suppression and clustered OFDM based anti-jamming modulation. A polynomial-based parameter estimator is proposed to provide parameter estimation for MMSE-DC. An

adaptive algorithm for polynomial-based estimator is developed. For military communications, a clustered OFDM based anti-jamming modulation is proposed. The proposed scheme with channel coding has great anti-jamming capability. In Chapter 4, we investigate MIMO-OFDM for wirelss LANs. An MMSE channel estimator is developed. The estimator can fully exploit the spatial and frequency correlations among MIMO-OFDM channels. The optimal training sequences are investigated so that the estimation error is minimized. We also investigate the linear precoding for MIMO-OFDM with limited feedback. A subspace tracking based approach is proposed to reduce the feedback rate.

CHAPTER II

INTER-CARRIER INTERFERENCE SUPPRESSION

As a promising technique for high-data-rate transmission, OFDM has been successfully used in many environments. However, time variations of wireless channels over one OFDM symbol period destroy orthogonality among subchannels and cause the ICI. Several methods [23]-[32] have been proposed to reduce the effect of ICI. One commonly used method is frequency domain equalization [23], [32]. In [23], a pilot symbol assisted frequency domain equalizer was proposed. In [24], an equalization technique suitable to time-varying multipath channels was developed. Antenna diversity is an effective way to combat the fading effect of wireless channels and can reduce the ICI, as shown in [16]. Another way to deal with the ICI is time domain windowing [25], [26]. In [27], ICI suppression for MIMO-OFDM was studied. Recently, a *self ICI cancellation* approach [28] was proposed, which transmits each symbol over a pair of adjacent subchannels with a 180^{o} phase shift. This method can suppress ICI significantly with reduced bandwidth efficiency. *Partial response coding* (PRC) in the time domain was studied for single-carrier systems to reduce the sensitivity to time offset [1] without sacrificing bandwidth efficiency. In the frequency domain, the PRC with correlation polynomial $F(D) = 1 - D$ was used to mitigate the ICI caused by carrier frequency offset in [29]. In this chapter, a general frequency domain PRC is proposed to suppress the ICI caused by Doppler frequency shift or carrier frequency offset.

In Section 2.1, an OFDM system with PRC is described and the exact ICI expression is derived. The optimal weights of PRC that minimizes the ICI power are obtained and the performance of the PRC with the optimal weights is analyzed in Section 2.2. Finally, the numerical and simulation results are presented in Section 2.3.

2.1 ICI for OFDM with PRC

As in [30], the baseband model of an OFDM system with PRC is shown in Figure 2.1. At the transmitter, the modulated signal is encoded by PRC. Let x_k's be the

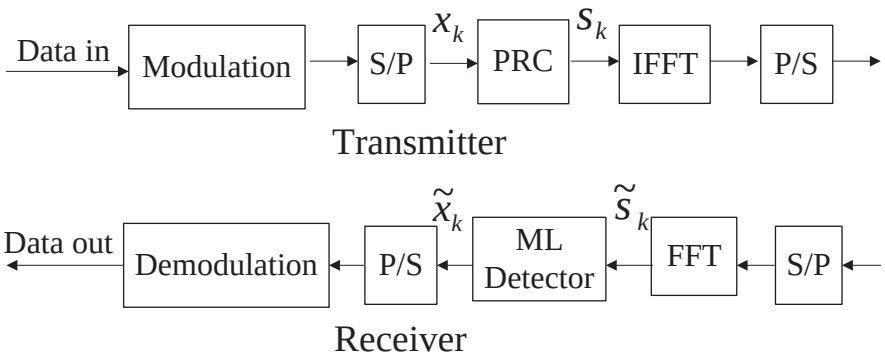


Figure 2.1: An OFDM system with PRC.

symbols to be transmitted and c_i's be the weights of PRC with unit norm, i.e.,

$$\sum_{i=0}^{K-1} c_i^2 = 1,$$

where K is the number of weights of PRC. Without loss of generality, we assume $E|x_k|^2 = 1$ and $E(x_k x_l^*) = 0$ for $k \neq l$. Then, the transmitted signal at the k-th subchannel can be expressed as

$$s_k = \sum_{i=0}^{K-1} c_i x_{k-i}. \tag{2.1}$$

The coded signal can be recovered by an ML sequence detector [33] at the receiver. The OFDM signal in the time domain is

$$y(t) = \sum_k s_k e^{\jmath 2\pi f_k t}, \quad 0 \le t < T_s, \tag{2.2}$$

where $f_k = f_0 + k\triangle f$ is the frequency of the k-th subchannel, $\triangle f = 1/T_s$ is the subchannel spacing, and T_s is the symbol duration. After passing through a time-varying channel with impulse response, $h(t,\tau)$, the received signal is

$$\tilde{y}(t) = \int h(t,\tau) y(t-\tau) \mathrm{d}\tau. \tag{2.3}$$

The channel impulse response for the frequency selective fading channel can be described as

$$h(t,\tau) = \sum_l \gamma_l(t)\delta(\tau - \tau_l), \tag{2.4}$$

where τ_l is the delay of the l-th path and $\gamma_l(t)$ is the corresponding path gain. Here, we assume that the complex stochastic processes $\gamma_l(t)$'s are independent for different l's and have the same statistics but different variance ε_l. For simplicity, we first consider the flat fading channel and omit the subscription l. Then, the received signal becomes

$$\tilde{y}(t) = \gamma(t) y(t).$$

The demodulated signal can be written as

$$\tilde{s}_m = \frac{1}{T_s} \int_0^{T_s} \tilde{y}(t) e^{-\jmath 2\pi f_m t} \mathrm{d}t. \tag{2.5}$$

Here, the integration is used instead of DFT. As indicated in [19], the difference is negligible. It was derived in [19] that the demodulated signal can be expressed as

$$\tilde{s}_m = a_0 s_m + \sum_{k \ne m} a_{m-k} s_k, \tag{2.6}$$

where a_l is defined as

$$a_l = \frac{1}{T_s} \int_0^{T_s} \gamma(t) e^{-\jmath 2\pi l \triangle f t} \mathrm{d}t. \tag{2.7}$$

In the above equation (2.5), a_0 is the gain of the desired signal and a_l's, for $l \neq 0$, represent the gains of the interfering signals from other subchannels. For time-invariant channels, $\gamma(t)$ is a constant and $a_l = 0$ for $l \neq 0$; consequently, there is no ICI. In general, for time-varying channels, $a_l \neq 0$ for some $l \neq 0$, the ICI exists.

The total ICI power is defined as

$$P_{\text{ICI}} = E\left|\sum_{l\neq 0} a_l s_{m-l}\right|^2. \tag{2.8}$$

For OFDM without PRC [19], it is

$$\tilde{P}_{\text{ICI}} = 1 - 2\int_0^{f_d} P(f)\text{sinc}^2(fT_s)\text{d}f, \tag{2.9}$$

where f_d is the maximum Doppler frequency shift, $P(f)$ is the power spectral density of $\gamma(t)$, and $\text{sinc}(x) = \sin(\pi x)/(\pi x)$.

It is derived in the appendix that P_{ICI} for OFDM with PRC can be expressed as

$$P_{\text{ICI}} = 1 - 2\int_0^{f_d} P(f)\text{sinc}^2(fT_s)\text{d}f + I_{\text{PRC}}(\mathbf{c}_K, f_dT_s), \tag{2.10}$$

where

$$\begin{aligned} I_{\text{PRC}}(\mathbf{c}_K, f_dT_s) &= \int_0^{f_d} \frac{8\sin(\pi fT_s)^2 P(f)}{\pi^2}\left(\sum_{k=1}^{K-1}\sum_{i=0}^{K-1-k}\frac{c_ic_{i+k}}{k^2 - f^2T_s^2}\right)\text{d}f \\ &= \int_0^{f_d} \frac{8\sin(\pi fT_s)^2 P(f)}{\pi^2} I(\mathbf{c}_K, fT_s)\text{d}f, \end{aligned} \tag{2.11}$$

with

$$\mathbf{c}_K = [c_0, c_1, \cdots\cdots, c_{K-1}]^T,$$

and

$$I(\mathbf{c}_K, fT_s) = \sum_{k=1}^{K-1}\sum_{i=0}^{K-1-k}\frac{c_ic_{i+k}}{k^2 - f^2T_s^2}.$$

2.2 Optimum PRC for OFDM

In the previous section, we introduced an OFDM system with PRC and derived the expression of the ICI power. In this section, we investigate the optimal PRC weights and analyze the corresponding performance.

2.2.1 Optimal Weights for PRC

From (2.10), the ICI power includes two parts: the ICI power for OFDM without PRC, and $I_{\mathrm{PRC}}(\mathbf{c}_K, f_d T_s)$ contributed by PRC. Therefore, the only way to reduce the ICI power is to minimize $I_{\mathrm{PRC}}(\mathbf{c}_K, f_d T_s)$ with respect to $\mathbf{c}_K$. In the integral of $I_{\mathrm{PRC}}(\mathbf{c}_K, f_d T_s)$, the first part, $(8\sin(\pi f T_s)^2 P(f)/\pi^2)$, is always positive. Therefore, we need only make the last part as small as possible. When $f^2 T_s^2 \ll 1$, the last part can be approximated as

$$I(\mathbf{c}_K, fT_s) \approx g(\mathbf{c}_K) = \sum_{k=1}^{K-1} \sum_{i=0}^{K-1-k} \frac{c_i c_{i+k}}{k^2} = \mathbf{c}_K^T \mathbf{R}_K \mathbf{c}_K, \tag{2.12}$$

where $\mathbf{R}_K$ is defined as

$$\mathbf{R}_K = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{8} & \cdots & \frac{1}{2(K-1)^2} \\ \frac{1}{2} & 0 & \frac{1}{2} & \cdots & \frac{1}{2(K-2)^2} \\ \frac{1}{8} & \frac{1}{2} & 0 & \cdots & \frac{1}{2(K-3)^2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2(K-1)^2} & \frac{1}{2(K-2)^2} & \frac{1}{2(K-3)^2} & \cdots & 0 \end{bmatrix}.$$

From (2.12), it is clear that the optimal $\mathbf{c}_K$ that minimizes the ICI is the normalized eigenvector of $\mathbf{R}_K$ corresponding to the smallest eigenvalue. Then,

$$I_{\min}(\mathbf{c}_K, fT_s) \approx g_{\min}(\mathbf{c}_K) = \lambda_0^{(K)},$$

where $\lambda_0^{(K)} \le \lambda_1^{(K)} \le \cdots \le \lambda_{K-1}^{(K)}$ are the ordered eigenvalues of $\mathbf{R}_K$.

When $K \to \infty$, we can obtain the limit of $g_{\min}(\mathbf{c}_K)$ using Corollary 4.2 in [35],

$$\lim_{K\to\infty} g_{\min}(\mathbf{c}_K) = \min_{\lambda\in\Re} f(\lambda), \tag{2.13}$$

where $f(\lambda)$ is defined as

$$f(\lambda) = \sum_{k\neq 0} \frac{1}{2k^2} e^{\jmath k\lambda}. \tag{2.14}$$

The minimum of $f(\lambda)$ can be found by setting its derivative with respect to λ to be zero, that is,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} f(\lambda) = -\sum_{k=1}^{\infty} \frac{1}{k} \sin(k\lambda) = 0. \tag{2.15}$$

Solving (2.15) for λ and substituting it into (2.13) leads to

$$\lim_{K\to\infty} g_{\min}(\mathbf{c}_K) = \sum_{k=1}^{\infty} \frac{(-1)^k}{k^2} = -\frac{\pi^2}{12}. \tag{2.16}$$

Once $g_{\min}(\mathbf{c}_K)$ is found, the total ICI power for OFDM with optimum PRC can be calculated using (2.10).

2.2.2 Performance Analysis

The exact expression of the ICI is too complicated to provide much insight. In many cases, it is difficult to calculate the ICI because the exact power spectrum is not available. Here we derive an approximate expression using a similar method in [19].

If $fT_s \ll 1$, we have the following approximation:

$$\sin^2(\pi f T_s) \approx (\pi f T_s)^2. \tag{2.17}$$

Substituting (2.12) and (2.17) into (2.11), we can obtain the following expression:

$$\begin{aligned} I_{\mathrm{PRC}}(\mathbf{c}_K, f_d T_s) &\approx \int_0^{f_d} \frac{8(\pi f T_s)^2}{\pi^2} P(f) g(\mathbf{c}_K) \mathrm{d}f \\ &= 4\alpha g(\mathbf{c}_K)(f_d T_s)^2, \end{aligned} \tag{2.18}$$

where α is defined as

$$\alpha = \frac{2}{T_s^2} \int_0^{f_d} f^2 P(f) \mathrm{d}f, \tag{2.19}$$

which is dependent on the spectral density of $\gamma(t)$. It is calculated in [19] that $\alpha = 1/2$ for the classical Doppler spectrum (Jakes' model) and $\alpha = 1$ for OFDM with carrier offset.

Using (3.10) in [19] and (2.18), we can obtain an approximate expression of the ICI for PRC as

$$P_{\mathrm{ICI}} \approx \left[\frac{\pi^2}{3} + 4g(\mathbf{c}_K)\right] \alpha (f_d T_s)^2 \tag{2.20}$$

Table 1: Table of optimal weights

K	$\mathbf{c}_K$	$-g(\mathrm{c}_K)$	Gain(dB)
1	1	0	0
2	0.7071, -0.7071	0.5000	4.066
3	-0.4775, 0.7376, -0.4775	0.6474	6.719
4	-0.3501, 0.6144, -0.6144, 0.3501	0.7137	8.786
∞		0.8225	———

The expression is much simpler than the exact one. From (2.18), the total ICI is decided by three factors: α, f_dT_s, and $g(\mathbf{c}_K)$. f_dT_s can be regarded as the normalized Doppler frequency shift with respect to the sunchannel spacing, Δf. $g(\mathbf{c}_K)$ represents the ICI reduction resulting from PRC. For OFDM without PRC [19],

$$\tilde{P}_{\mathrm{ICI}} \approx \frac{\pi^2}{3}\alpha(f_dT_s)^2. \tag{2.21}$$

Then, the performance gain of PRC is

$$\begin{aligned} \mathrm{Gain(dB)} &= \tilde{P}_{\mathrm{ICI}}(\mathrm{dB}) - P_{\mathrm{ICI}}(\mathrm{dB}) \\ &\approx -10\log_{10}\left[1+\frac{12}{\pi^2}g(\mathbf{c}_K)\right]. \end{aligned}$$

Table 1 lists the optimal weights, the corresponding $g(\mathbf{c}_K)$, and the performance gain according to the above results. From the table, the value of $g(\mathbf{c}_K)$ is close to the limit when $K = 4$.

In the above discussions, we obtained the optimum PRC and analyzed its performance for flat fading channels. For frequency selective fading channels, it is usually assumed that a cyclic extension is inserted ahead of each OFDM symbol to combat the ISI. Then, the expression of the ICI for frequency selective fading channels can be derived in a similar way

$$P_{\mathrm{ICI}} = 1 - 2\int_0^{f_d} P(f)\mathrm{sinc}^2(fT_s)\mathrm{d}f + \tilde{I}_{\mathrm{PRC}}(\mathbf{c}_K, f_dT_s), \tag{2.22}$$

where

$$\tilde{I}_{\mathrm{PRC}}(\mathbf{c}_K, f_dT_s) = \int_0^{f_d} \frac{8\sin(\pi fT_s)^2 P(f)}{\pi^2}\tilde{I}(\mathbf{c}_K, fT_s)\mathrm{d}f, \tag{2.23}$$

and

$$\tilde{I}(\mathbf{c}_K, fT_s) = \sum_{k=1}^{K-1} \sum_{i=0}^{K-1-k} \sum_{l} \varepsilon_l e^{\jmath 2\pi k \Delta f \tau_l} \frac{c_i c_{i+k}}{k^2 - f^2 T_s^2}. \tag{2.24}$$

Usually, the path delay τ_l is much smaller than the symbol duration. Then, for small size PRC, the term $\varepsilon_l e^{\jmath 2\pi k \Delta f \tau_l}$ is approximately a constant. And the optimum PRC with small size obtained for flat fading channels is also applicable to the frequency selective fading channels.

2.3 Numerical and Simulation Results

In this section, we present the numerical and simulation results to show the performance improvement of OFDM with the optimum PRC.

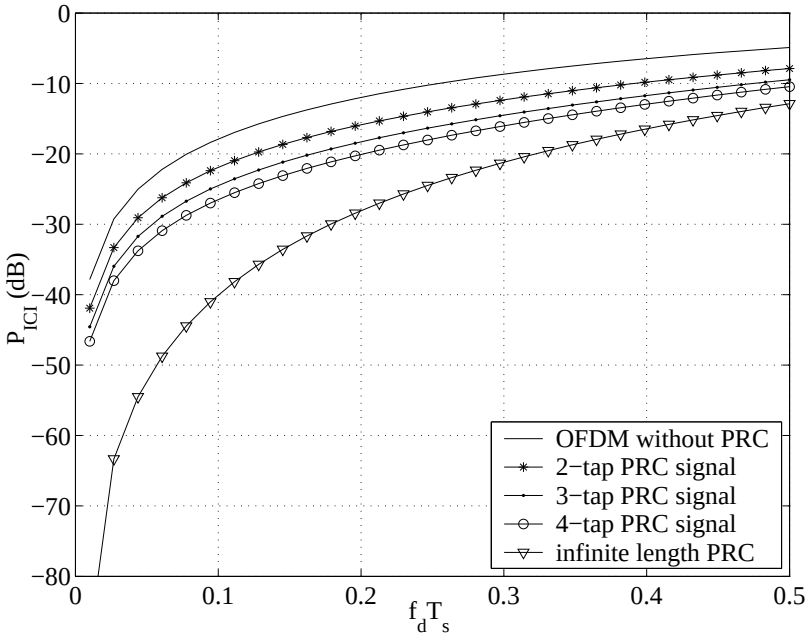


Figure 2.2: Comparison of the ICI power due to Doppler frequency shift.

2.3.1 Numerical Results

To examine the performance improvement, we compare the ICI value of OFDM with and without PRC.

Figure 2.2 presents a comparison of the ICI power for the classical Doppler spectrum (Jakes' model, [8]). From the figure, the ICI power is reduced by about 4.0 dB for a two-tap and about 6.2 dB for a three-tap PRC, respectively. The optimum PRC can also reduce the ICI resulting from carrier offset, as shown in Figure 2.3. There is about a 4.5 dB improvement for a two-tap PRC.

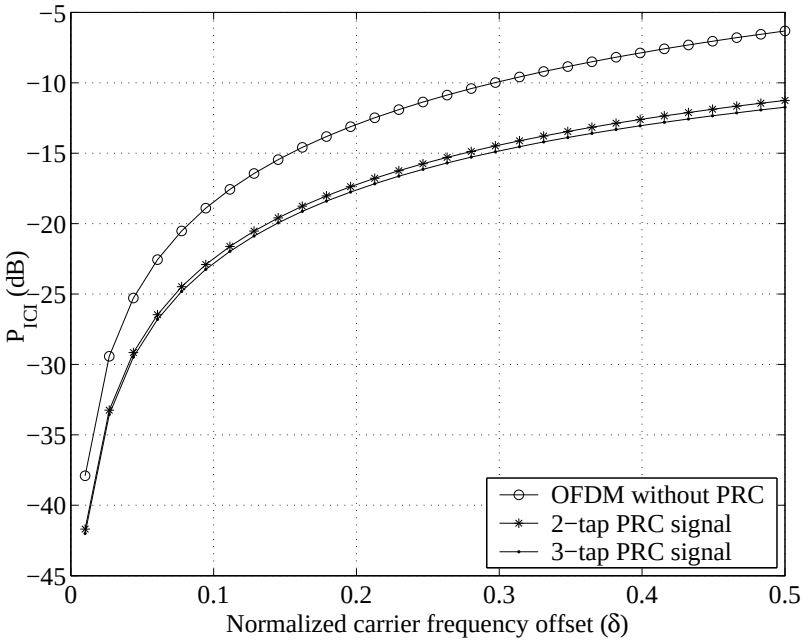


Figure 2.3: Comparison of the ICI power due to carrier offset.

2.3.2 Simulation Results

Here, we present our simulation results for the performance of PRC. The system model and parameters used in our simulation are the same as those in [38]. The entire channel bandwidth (800 kHz) is divided into 128 subchannels. The 120 subchannels at the middle are used to transmit data. The remaining subchannels on each side are used as guard subchannels. QPSK with coherent demodulation is used. A (40,20) R-S code, with each code symbol consisting of three QPSK symbols grouped in frequency, is used so that each block forms an R-S codeword. The noise is assumed to be white Gaussian with zero-mean and variance σ_n^2. Then, the SNR is $1/\sigma_n^2$. The time-varying

fading channel is generated using Jakes' model. A cyclic prefix is used to avoid ISI.

Figure 2.4 shows *word-error-rate* (WER) versus SNR for *hilly terrain* (HT) chan-

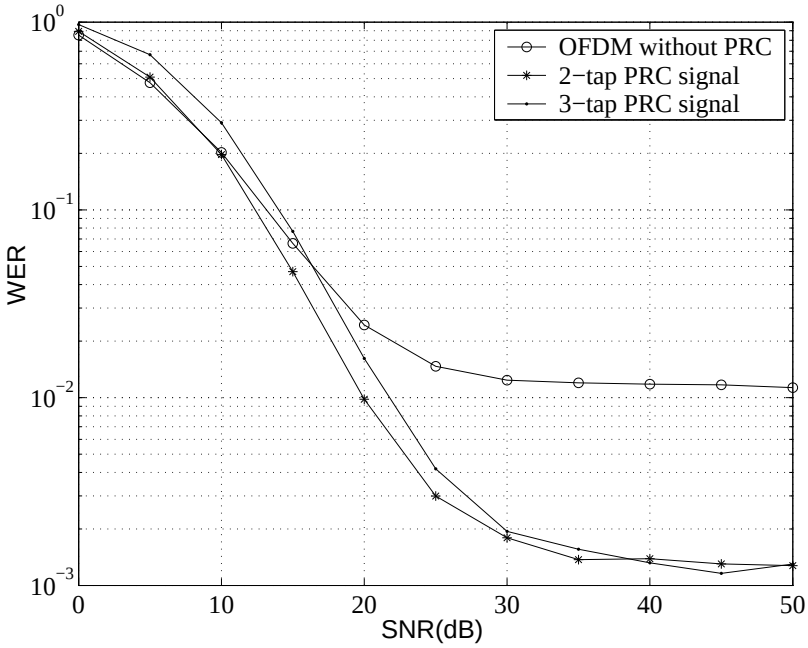


Figure 2.4: WER of OFDM with and without PRC ($f_d T_s = 0.1$).

nel. From the figure, OFDM with PRC has some performance loss at low SNR. However, for the system at high SNR, the ICI is the dominant impairment and OFDM with PRC has better performance than that without PRC. The error floor resulting from Doppler frequency shift is reduced from 10^{-2} to 10^{-3}. Because PRC has some performance loss when $K > 2$ as for time domain PRC [1], the performance of a three-tap PRC is not as good as that of a two-tap PRC, though the three-tap PRC has better performance in terms of ICI suppression.

The error floor versus normalized Doppler frequency shift is shown in Figure 2.5. From the figure, to ensure WER below 1%, the maximum tolerable Doppler frequency shift for OFDM without PRC is about 10% of the subchannel spacing, and it is relaxed to be 15% for OFDM with a two-tap or a three-tap PRC.

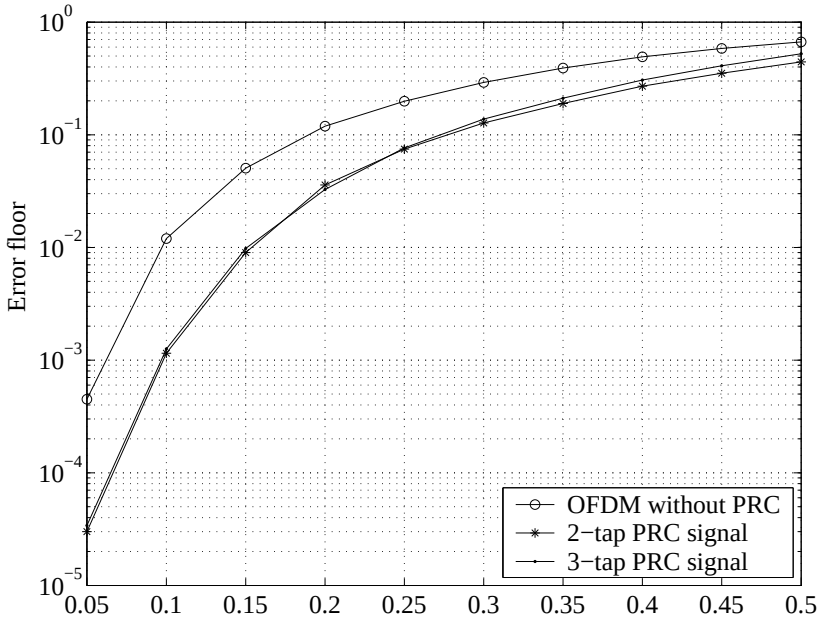


Figure 2.5: Error floor comparison for OFDM signal with and without PRC.

2.4 Summary

In this chapter, we proposed a frequency domain PRC to reduce the ICI caused by the time variations of wireless channels. The optimal weights for PRC that minimize the ICI power are obtained. The numerical and simulation results show that PRC effectively reduces the error floor caused by Doppler frequency shift or carrier offset. Even though our discussions are for flat fading channels, the obtained optimal PRC weights are also applicable to the frequency selective channels.

2.5 Appendix: Derivation of the ICI Power

Here, we derive the ICI power caused by Doppler frequency shift for OFDM with PRC. First, we calculate the autocorrelation of a_l:

$$E(a_m a_n^*) = E\left\{\frac{1}{T_s}\int_0^{T_s}\gamma(t)e^{-\jmath 2\pi m\Delta f t}\mathrm{d}t \times \frac{1}{T_s}\int_0^{T_s}\gamma^*(\tau)e^{-\jmath 2\pi n\Delta f \tau}\mathrm{d}\tau\right\}, \quad (2.A.1)$$

where $m \neq n$. By (2.7) and Equation (10-50) in [36],

$$E(a_m a_n^*) = \int_{-1}^{1} r(T_s x)\frac{-\sin[\pi(m-n)|x|]}{\pi(m-n)}e^{-\jmath 2\pi\frac{m+n}{2}x}\mathrm{d}x, \quad (2.A.2)$$

where $r(\tau)$ is the autocorrelation of $\gamma(t)$. Then, the ICI power is

$$\begin{aligned} P_{\text{ICI}} &= E\left|\sum_{l\neq 0} a_l s_{m-l}\right|^2 \\ &= E\left|\sum_{l\neq 0} a_l \sum_{i=0}^{K-1} c_i x_{m-l-i}\right|^2 \\ &= \sum_{n\neq i}\sum_{i=0}^{K-1} c_i^2 E|a_{n-i}|^2 + 2\sum_{\substack{n\neq i\\ n\neq i+k}}\sum_{k=0}^{K-1}\sum_{i=0}^{K-1-k} c_i c_{i+k} E(a_{n-i}a_{n-i-k}^*) \\ &= \sum_{l\neq 0} E|a_l|^2 + 2\sum_{k=1}^{K-1}\sum_{i=0}^{K-1-k} c_i c_{i+k} T_k, \end{aligned}$$

where we use the identity $\sum_{i=0}^{K-1} c_i^2 = 1$ and definition

$$T_k = \sum_n E(a_{n-i}a_{n-i-k}^*) - E(a_0 a_{-k}^*) - E(a_k a_0^*). \tag{2.A.3}$$

Substituting (2.A.2) into (2.A.3), we have

$$T_k = \int_{-1}^{1} r(T_s x)\frac{-\sin(\pi k|x|)}{\pi x} \times \left[\sum_n e^{-\jmath 2\pi \frac{2(n-i)-k}{2}x} - 2\cos(\pi k x)\right]\mathrm{d}x \tag{2.A.4}$$

From Equation (3.2) and (3.5) in [37],

$$\sum_{n=-\infty}^{\infty} e^{-\jmath 2\pi n x} = \sum_{m=-\infty}^{\infty} \delta(x-m).$$

Consequently,

$$\begin{aligned} T_k &= \int_{-1}^{1} r(T_s x)\frac{-\sin(\pi k|x|)}{\pi k}\left[\sum_m \delta(x-m)e^{\jmath\pi k x} - 2\cos(\pi k x)\right]\mathrm{d}x \\ &= \int_{-1}^{1} r(T_s x)\frac{2\sin(\pi k|x|)\cos(\pi k x)}{\pi k}\mathrm{d}x \\ &= \int_{-1}^{1} \frac{\sin(2\pi k|x|)}{\pi k}\left[2\int_0^{f_d} P(f)\cos(2\pi f T_S x)\mathrm{d}f\right]\mathrm{d}x \\ &= \int_0^{f_d} \frac{4\sin^2(\pi f T_s)P(f)}{\pi^2(k^2 - f^2T_s^2)}\mathrm{d}f. \end{aligned} \tag{2.A.5}$$

From [19],

$$\sum_{l\neq 0} E|a_l|^2 = 1 - 2\int_0^{f_d} P(f)\mathrm{sinc}^2(fT_s)\mathrm{d}f, \tag{2.A.6}$$

then

$$P_{\mathrm{ICI}} = 1 - 2\int_0^{f_d} P(f)\mathrm{sinc}^2(fT_s)\mathrm{d}f + I_{\mathrm{PRC}}(c_K, f_dT_s), \tag{2.A.7}$$

where

$$I_{\mathrm{PRC}}(\mathbf{c}_K, f_dT_s) = \int_0^{f_d} \frac{8\sin(\pi fT_s)^2 P(f)}{\pi^2} \left(\sum_{k=1}^{K-1} \sum_{i=0}^{K-1-k} \frac{c_i c_{i+k}}{k^2 - f^2T_s^2} \right) \mathrm{d}f.$$

CHAPTER III

CLUSTERED OFDM AND ITS MILITARY APPLICATIONS

This chapter introduces a novel system, called clustered OFDM. We investigates interference suppression and military applications of clustered OFDM systems. Section 3.1 introduces the concept of clustered OFDM. In Section 3.2, we propose a polynomial-based parameter estimator for clustered OFDM with adaptive antenna arrays for interference suppression. In Section 3.3, a clustered OFDM based anti-jamming modulation is proposed for military communications.

3.1 Clustered OFDM

Recently, clustered OFDM [20]-[22] was proposed for high-speed wireless transmission [39]. In a clustered OFDM system, contiguous OFDM subchannels are grouped into subchannel clusters and each user accesses several clusters at different frequencies. Figure 3.1 shows the concept of clustered OFDM [46]. A wideband OFDM signal is

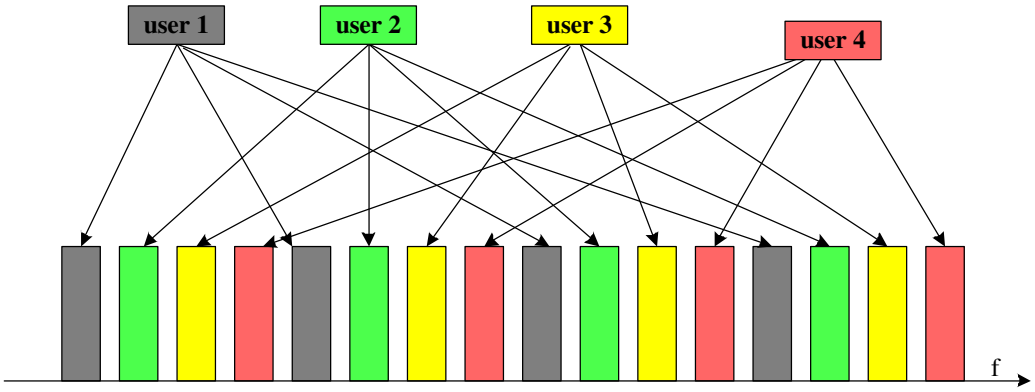


Figure 3.1: Concept of clustered OFDM.

divided into 16 clusters. Four users utilize all clusters and each accesses four clusters. Since the clusters belonging to one user are distributed over the whole frequency band, frequency diversity can be achieved for frequency selective fading channels by means of channel coding [20]-[22]. Since the size of each cluster in a clustered OFDM system is smaller than the whole bandwidth, the PAPR can be reduced by $10 \log M$ if the whole band is divided into M clusters. In addition, the complexity of nonlinear coding for PAPR reduction can be significantly reduced [20]. Clustered OFDM can also be combined with multiple transmit antennas to achieve transmit diversity. In a clustered OFDM system with multiple transmit antennas, each cluster can be assigned to an independent antenna, a coding scheme is used to obtain the transmit diversity. If a feedback channel is available, the cluster can be assigned to the antenna with the best channel quality.

Clustered OFDM is also an ideal modulation for joint physical and medium access layer optimization [40]. The system has a high degree of flexibility for supportable bit rates since it can adaptively allocate different clusters to different users.

3.2 Interference Suppression for Clustered OFDM

For cellular wireless communication systems, one of the major limitations is co-channel interference. The co-channel interference arises when two or more users transmit signal simultaneously on the same channel. In TDMA systems, such as GSM/GPRS and IS-136, the power of co-channel interference mainly depends on the reuse factor, and cannot be avoided since the same channels must be reused in some other cells not far away. Various techniques have been developed to reduce the co-channel interference, for example, dynamic frequency allocation, power control, and adaptive multi-rate coding. Adaptive antenna arrays have been proven to be an effective technique to mitigate fading effect and suppress co-channel interference [42]. In [41], MMSE diversity combing is proposed to suppress co-channel interference.

Among many algorithms for calculating the weights for MMSE diversity combining, *direct matrix inversion with diagonal loading* (DMI/DL) [43], [44] is able to reach the optimal performance. However, this algorithm requires accurate information of the channel responses corresponding to the desired signals and the correlations of the received signals from different receive antennas.

For classical OFDM, either pilot-symbol aided or decision-directed channel estimators can be used to obtain the channel information. When the channel statistics are available, an MMSE estimation with the optimal transform obtained through decomposing the channel correlation matrix was proposed to achieve the optimal performance [63], [64]. However, extremely high computational complexity limits the practical use of the optimal transform. To overcome the difficulties of obtaining the optimal transform, a robust estimation using DFT to approximate the optimal transform was proposed in [38]. It was proven that the leakage resulting from the approximation is negligible if the number of subcarriers is large enough, which is usually satisfied for classical OFDM systems. A similar estimation approach was used to estimate the correlation of the received signals for OFDM systems with adaptive antenna arrays [45]. Unfortunately, when it is applied into clustered OFDM systems, the leakage is very large because of the small size of each cluster and the performance loss is unacceptable. In [22], some robust transforms for clustered OFDM channel estimation were proposed to reduce the edge effect. We propose to use a polynomial-based estimator to obtain the desired channel information and the received signal correlation for clustered OFDM systems. The motivation is that polynomial approximation can achieve high accuracy in a small area.

Polynomial approximation has been extensively studied in mathematical literature [65], [66]. Applications of the polynomial-based estimation to communication systems can also be found in [67], [68]. For the polynomial-based estimation, the channel is assumed to be a smooth function of time or frequency, and can be approximated by a

polynomial function. Two critical issues for a polynomial-based estimator have to be taken into account: the polynomial order and window size. With the increase of the window size, the approximation error increases while the variance of estimation error is reduced [67]. For a slow-varying channel, a fixed window size is good enough. When a channel changes fast, an adaptive window size is desirable to improve performance. In [67] and [68], a tentative search algorithm was used to find a local optimal window size. However, this method is sensitive to noise and requires high computational complexity.

In this section, we investigate polynomial-based estimation for clustered OFDM. We first study the impact of the polynomial order and window size on the estimation error. Then, we derive an approximately optimal window size and propose an adaptive algorithm with low complexity to obtain the optimal window size.

3.2.1 Clustered OFDM with Adaptive Antenna Arrays

As in [46], the baseband model of a clustered OFDM system with adaptive antenna arrays for interference suppression is shown in Figure 3.2. After OFDM demodulation,

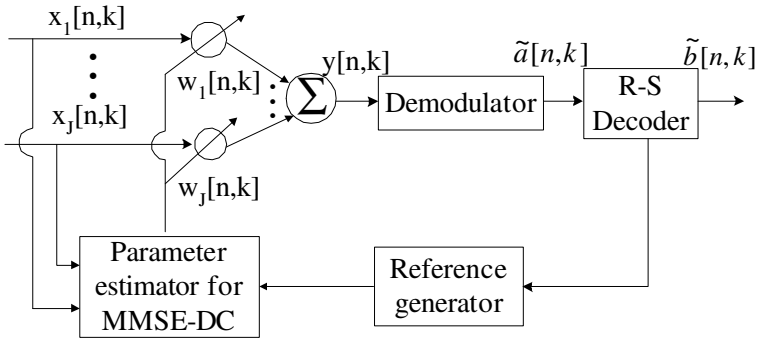


Figure 3.2: Clustered OFDM with the MMSE diversity combiner.

the received signals from different receive antennas at the same block and subcarrier are multiplied by different weights and the weighted signals are summed to form the output signals. The output signals are further demodulated and decoded. The weights are determined to maximize the output *signal-to-noise-plus-interference ratio*

(SINR). For a J-branch diversity system, the received signal from the m-th antenna at the k-th subcarrier of the n-th block can be written as

$$x_m[n,k] = \sum_{l=0}^{L} H_m^{(l)}[n,k]s_l[n,k] + n_m[n,k], \tag{3.1}$$

where $s_0[n,k]$ is the desired signal, $H_m^{(0)}[n,k]$ is the channel frequency response corresponding to the desired signals, $s_l[n,k]$ is the l-th interferer, $H_m^{(l)}[n,k]$ is the corresponding channel frequency response, and $n_m[n,k]$ is AWGN with zero-mean and variance σ_n^2.

In the above expression, we assume synchronized co-channel interferers. As indicated in [45], the effect of the asynchronous interference on the system performance is similar to the synchronous interference. We also assume the signals from the desired user and interferers are *i.i.d.* complex random variables with zero-mean and unit variance. And $H_m^{(l)}[n,k]$'s for different m's or l's are independent, stationary, and complex Gaussian with zero-mean and different variances σ_l^2's.

The received signals are linearly combined with weights, $w_m[n,k]$'s,

$$y[n,k] = \mathbf{w}^H[n,k]\mathbf{x}[n,k], \tag{3.2}$$

where $\mathbf{x}[n,k]$ is the received signal vector, defined by $(x_1[n,k], \cdots, x_J[n,k])^T$, and $\mathbf{w}[n,k]$ is the weight vector, defined by $(w_1[n,k], \cdots, w_J[n,k])^T$. The weight vector can be calculated by the DMI/DL algorithm [44]:

$$\mathbf{w}[n,k] = (\mathbf{R}[n,k] + \gamma\mathbf{I})^{-1}\mathbf{H}^{(0)}[n,k], \tag{3.3}$$

where γ is a diagonal loading factor [44], [45], $\mathbf{I}$ is $J \times J$ identity matrix, and $\mathbf{H}^{(0)}[n,k]$ is the $J \times 1$ channel response vector corresponding to the desired signal. $\mathbf{R}[n,k]$ is the $J \times J$ correlation matrix, defined as

$$\begin{aligned}\mathbf{R}[n,k] &= (r_{ij}[n,k])_{i,j=1}^{J} = \left(\mathrm{E}_c\left\{x_i[n,k]x_j^*[n,k]\right\}\right)_{i,j=1}^{J} \\ &= \left(\sum_{l=0}^{L} H_i^{(l)}[n,k]H_j^{(l)*}[n,k] + \sigma_n^2\delta[i-j]\right)_{i,j=1}^{J},\end{aligned} \tag{3.4}$$

where E_c is the conditional expectation given the channel parameters corresponding to both the desired signal and interferers.

From (3.3), to obtain the weight vector for adaptive antenna arrays, the instantaneous correlations of the signals from different antennas and channel parameters have to be estimated. The accuracy of estimation directly affects the performance of the system.

3.2.2 Polynomial-Base Parameter Estimation

As indicated before, to obtain the optimal weights for the MMSE combiner, we have to estimate channel parameters and instantaneous correlations. We use the polynomial-based estimator to obtain those parameters. Since the polynomial estimator is applicable to both channel and instantaneous correlation estimation, we focus only on estimation of the instantaneous correlations.

3.2.2.1 Polynomial model for instantaneous correlation estimation

It was shown in [38] that the correlation of the channel responses satisfies *separation property*, i.e., the correlation of the channel responses in time and frequency domain $r_H(\Delta t, \Delta f)$ can be expressed as the product of the time domain correlation $r_T(\Delta t)$ and frequency domain correlation $r_F(\Delta f)$. From [45], the separation property holds for instantaneous correlations of the received signals from different antennas. With this property, a two-dimensional polynomial fitting can be simplified to two one-dimensional fitting problems: a frequency domain fitting and a time domain fitting. As a result, the computational complexity is reduced significantly. Although the autocorrelations of the received signals do not satisfy the separation property, we still use two one-dimensional estimators for simplicity. The resultant performance is satisfactory, as shown later by our simulation. Furthermore, the real part and imaginary part of instantaneous correlations can be estimated independently. Therefore, we assume that all variables are real for the rest of the derivations.

The temporal estimation of the instantaneous correlations $\tilde{r}_{ij}[n,k]$ can be written as

$$\begin{aligned}\tilde{r}_{ij}[n,k] &= x_i[n,k]x_j^*[n,k] \\ &= r_{ij}[n,k] + \tilde{v}_{ij}[n,k], \end{aligned} \tag{3.5}$$

where $\tilde{v}_{ij}[n,k]$ is the temporal estimation error

$$\begin{aligned}\tilde{v}_{ij}[n,k] = & \sum_{l_1,l_2=0,l_1\neq l_2}^{L} H_i^{(l_1)}[n,k]H_j^{(l_2)*}[n,k]s_{l_1}[n,k]s_{l_2}^*[n,k] \\ & + n_i[n,k]\sum_{l_2=0}^{L} H_j^{(l_2)*}s_{l_2}^*[n,k] \\ & + n_j^*[n,k]\sum_{l_1=0}^{L} H_i^{(l_1)}s_{l_1}[n,k]. \end{aligned} \tag{3.6}$$

Since $s_l[n,k]$'s are *i.i.d*, $\tilde{v}_{ij}[n,k]$ is white with $E(\tilde{v}_{ij}[n,k]) = 0$, and $\mathrm{Var}(\tilde{v}_{ij}[n,k]) = \tilde{\sigma}_{ij}^2[n,k]$. For MMSE estimation, only second-order statistics, the variance $\tilde{\sigma}_{ij}^2[n,k]$ of $\tilde{v}_{ij}[n,k]$, is concerned. The channel parameters for the desired signal and interferers change with time and frequency, so is the variance of $\tilde{v}_{ij}[n,k]$. Since the estimator works independently for each pair of i and j, we eliminate the subscript i,j in the rest of the section.

According to approximation theory [70], the instantaneous correlations within a $(2N+1)$ point window centered at n_0 can be approximated by a polynomial of order M,

$$r[n] = \sum_{m=0}^{M} a_m(n-n_0)^m + \varepsilon_r[n]. \tag{3.7}$$

Here, we have ignored the frequency index k. It is obvious from (3.7) that $r[n_0] = a_0$. With the temporal estimation, the coefficients of the polynomial can be obtained by solving the following weighted least-square equation:

$$\min_{a_0,\cdots,a_M} \sum_{n=n_0-N}^{n_0+N} \left| \tilde{r}[n] - \sum_{m=0}^{M} a_m(n-n_0)^m \right|^2 W\left(\frac{n-n_0}{N}\right), \tag{3.8}$$

where $W\left(\frac{n-n_0}{N}\right)$ denotes a nonnegative weight function. Equation (3.8) can be written in matrix form

$$\min_{\mathbf{a}_M}(\tilde{\mathbf{r}} - \mathbf{D}_M\mathbf{a}_M)^T\mathbf{W}(\tilde{\mathbf{r}} - \mathbf{D}_M\mathbf{a}_M),$$

where

$$\mathbf{a}_M = (a_0, \cdots, a_M)^T,\ \tilde{\mathbf{r}} = (\tilde{r}[n_0 - N], \cdots, \tilde{r}[n_0 + N])^T,$$

$$\mathbf{W} = \text{diag}\{W(\frac{n_0 - N}{N}), \cdots, W(\frac{n_0 + N}{N})\},$$

and

$$\mathbf{D}_M = \begin{pmatrix} 1 & -N & \cdots & (-N)^M \\ 1 & 1-N & \cdots & (1-N)^M \\ \vdots & \vdots & \vdots & \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \\ 1 & N & \cdots & N^M \end{pmatrix}.$$

The estimation of $\mathbf{a}_M$ can be derived as

$$\hat{\mathbf{a}}_M = (\mathbf{D}_M^T\mathbf{W}\mathbf{D}_M)^{-1}\mathbf{D}_M^T\mathbf{W}\tilde{\mathbf{r}}. \tag{3.9}$$

Note that $\mathbf{D}_M^T\mathbf{W}\mathbf{D}_M$ is invertible provided that the window size N is larger than order M. Then, the estimation of $r[n_0]$ is $\hat{r}[n_0] = \mathbf{e}_1(\mathbf{D}_M^T\mathbf{W}\mathbf{D}_M)^{-1}\mathbf{D}_M^T\mathbf{W}\tilde{\mathbf{r}}$, where $\mathbf{e}_1 = (1, 0, \cdots, 0)$. In other words, the estimation of $r[n_0]$ is the first element of $\hat{\mathbf{a}}_M$.

The window size and polynomial order play important roles in polynomial-based estimation. Next, we investigate these issues.

Polynomial Order

To find the optimal polynomial order, we first study its impact on the estimation error. The MSE of estimation at the center point n_0 within $[n_0 - N, n_0 + N]$ is defined as

$$\varepsilon_p[n_0] = E\left\{|r[n_0] - \hat{r}[n_0]|^2\right\}, \tag{3.10}$$

where the expectation is over $\tilde{v}[n]$.

Let $\boldsymbol{\varepsilon}_p = E\{(\hat{\mathbf{a}}_M - \mathbf{a}_M)(\hat{\mathbf{a}}_M - \mathbf{a}_M)^T\}$. Then, direct calculation yields

$$\boldsymbol{\varepsilon}_p = \mathrm{Var}(\hat{\mathbf{a}}_M) + (\mathrm{bias}(\hat{\mathbf{a}}_M))\,(\mathrm{bias}(\hat{\mathbf{a}}_M))^T, \tag{3.11}$$

where $\mathrm{Var}(\hat{\mathbf{a}}_M) = E\{(\hat{\mathbf{a}}_M - E\{\hat{\mathbf{a}}_M\})(\hat{\mathbf{a}}_M - E\{\hat{\mathbf{a}}_M\})^T\}$, $\varepsilon_p[n_0]$ is the element at the first row and first column of $\boldsymbol{\varepsilon}_p$, and the bias is defined as

$$\mathrm{bias}(\hat{\mathbf{a}}_M) = E\{\hat{\mathbf{a}}_M\} - \mathbf{a}_M.$$

To find $\varepsilon_p[n_0]$, we need to evaluate the bias and variance of the coefficient estimation.

Define $\mathbf{G}_M = \mathbf{D}_M^T\mathbf{W}\mathbf{D}_M$, $\bar{\mathbf{G}}_M = \mathbf{D}_M^T\mathbf{W}^2\mathbf{D}_M$, and the model error $\boldsymbol{\xi}_M = \mathbf{r} - \mathbf{D}_M\mathbf{a}_M$, where $\mathbf{r}$ is the exact instantaneous correlation vector. From (3.9), we have,

$$\begin{aligned} E\{\hat{\mathbf{a}}_M\} &= E\{\mathbf{G}_M^{-1}\mathbf{D}_M^T\mathbf{W}\tilde{\mathbf{r}}\} \\ &= \mathbf{G}_M^{-1}\mathbf{D}_M^T\mathbf{W}\mathbf{r} \\ &= \mathbf{G}_M^{-1}\mathbf{D}_M^T\mathbf{W}(\mathbf{D}_M\mathbf{a}_M + \boldsymbol{\xi}_M) \\ &= \mathbf{a}_M + \mathbf{G}_M^{-1}\mathbf{D}_M^T\mathbf{W}\boldsymbol{\xi}_M. \end{aligned} \tag{3.12}$$

Then,

$$\mathrm{bias}(\hat{\mathbf{a}}_M) = \mathbf{G}_M^{-1}\mathbf{D}_M^T\mathbf{W}\boldsymbol{\xi}_M. \tag{3.13}$$

The estimation variance is

$$\begin{aligned} \mathrm{Var}(\hat{\mathbf{a}}_M) &= E\left\{(\mathbf{G}_M^{-1}\mathbf{D}_M^T\mathbf{W}\tilde{\mathbf{V}})(\mathbf{G}_M^{-1}\mathbf{D}_M^T\mathbf{W}\tilde{\mathbf{V}})^T\right\} \\ &= \tilde{\sigma}^2[n_0]\mathbf{G}_M^{-1}\bar{\mathbf{G}}_M\mathbf{G}_M^{-1}, \end{aligned} \tag{3.14}$$

where $\tilde{\mathbf{V}} = (\tilde{v}[n_0 - N], \cdots, \tilde{v}[n_0 + N])^T$ is the temporal estimation error vector, which is assumed to have constant variance $\tilde{\sigma}^2[n_0]$ within $(n_0 - N, n_0 + N)$.

Let $g_{ij}^{(M)}$ denote the element of the matrix $\mathbf{G}_M$ at the i-th row and j-th column.

From the definition of $\mathbf{G}_M$, we have

$$\begin{aligned} g_{ij}^{(M)} &= \sum_{n=-N}^{N} n^{i+j} W(\frac{n}{N}) \\ &= N^{i+j+1} \sum_{n=-N}^{N} \Big(\frac{n}{N}\Big)^{i+j} \frac{1}{N} W(\frac{n}{N}) \\ &\approx N^{i+j+1} \psi_{i+j}, \end{aligned} \tag{3.15}$$

where

$$\psi_p = \int_{-1}^{1} t^p W(t) \mathrm{d}t.$$

Define $\bar{\mathbf{N}}_M = \mathrm{diag}\{1, N, \cdots, N^M\}$ and $\mathbf{\Psi}_M = \{\psi_{i+j}\}_{0 \le i,j \le M}$, then

$$\mathbf{G}_M \approx N \bar{\mathbf{N}}_M \mathbf{\Psi}_M \bar{\mathbf{N}}_M. \tag{3.16}$$

Similarly, we have

$$\bar{\mathbf{G}}_M \approx N \bar{\mathbf{N}}_M \bar{\mathbf{\Psi}}_M \bar{\mathbf{N}}_M, \tag{3.17}$$

where $\bar{\mathbf{\Psi}}_M = \{\bar{\psi}_{i+j}\}_{0 \le i,j \le M}$ with $\bar{\psi}_p = \int_{-1}^{1} t^p W^2(t) \mathrm{d}t$. Note that ψ_p and $\bar{\psi}_p$ are zero for odd p when the weight function is symmetric. Then, the estimation variance becomes

$$\mathrm{Var}(\hat{\mathbf{a}}_M) \approx \frac{\tilde{\sigma}^2[n_0]}{N} \bar{\mathbf{N}}^{-1} \mathbf{\Psi}_M^{-1} \bar{\mathbf{\Psi}}_M \mathbf{\Psi}_M^{-1} \bar{\mathbf{N}}_M^{-1}. \tag{3.18}$$

According to Weierstrass Approximation Theorem [70], the remainder $\boldsymbol{\xi}_M$ can be approximated by a polynomial with any degree accuracy, i.e.,

$$\boldsymbol{\xi}_M = \sum_{k=M+1}^{\infty} a_k \mathbf{d}_k, \tag{3.19}$$

where $\mathbf{d}_k = \big((-N)^k, \cdots, N^k\big)^T$. From (3.13), (3.15), (3.16), and (3.19), the bias can be written as

$$\mathrm{bias}(\hat{\mathbf{a}}_M) \approx \sum_{k=M+1}^{\infty} N^k a_k \bar{\mathbf{N}}_M^{-1} \mathbf{\Psi}_M^{-1} \mathbf{\Phi}_k, \tag{3.20}$$

where $\mathbf{\Phi}_k = (\psi_k, \cdots, \psi_{k+M})^T$. The first element of $\mathbf{\Phi}_k$, ψ_k, is equal to zero for odd k and symmetric weight function. Consequently, the expressions of the approximate variance and bias of $\hat{r}[n_0]$ are given as follows,

$$\mathrm{Var}(\hat{r}[n_0]) \approx \frac{\tilde{\sigma}^2[n_0]}{N} \mathbf{e}_1 \bar{\mathbf{N}}_M^{-1} \mathbf{\Psi}_M^{-1} \bar{\mathbf{\Psi}}_M \mathbf{\Psi}_M^{-1} \bar{\mathbf{N}}_M^{-1} \mathbf{e}_1^T, \tag{3.21}$$

$$\mathrm{bias}(\hat{r}[n_0]) \approx \begin{cases} \sum_{k=M+1}^{\infty} N^k a_k \mathbf{e}_1 \bar{\mathbf{N}}_M^{-1} \mathbf{\Psi}_M^{-1} \mathbf{\Phi}_k, & M \text{ is odd}, \\ \sum_{k=M+2}^{\infty} N^k a_k \mathbf{e}_1 \bar{\mathbf{N}}_M^{-1} \mathbf{\Psi}_M^{-1} \mathbf{\Phi}_k, & M \text{ is even}. \end{cases} \tag{3.22}$$

From (3.22), for even M, the bias of the M-th order and $(M+1)$-th order polynomial-based estimators are related to at least the $(M+2)$-th coefficient of the polynomial. Then, the following property of the polynomial-based estimation is concluded:

For the even number M, the MSEs of estimation at the center with the M-th order and $M+1$-th order polynomial-based estimator are the same.

The proof is given in the appendix. We should note that the statement is valid only for estimation at the center of the window and symmetric weight function. Otherwise, it does not hold, as shown in [68]. The property can be further confirmed by simulation results in Figure 3.3.

In the simulation of Figure 3.3, the time domain correlation estimation is performed using polynomials with different orders and window sizes. From this figure, we have the following observations:

1. The *normalized MSE* (NMSE) performance is related to the window size, and an improper window size results in performance degradation.

2. For an even number M, the polynomial-based estimators with M-th order and $M+1$-th order have the same performance.

3. The estimation performance is not sensitive to the polynomial orders. For example, from zero-th order to second order, there is about a 1 dB improvement. However, there is only a 0.3 dB gain from second order to sixth order. Thus,

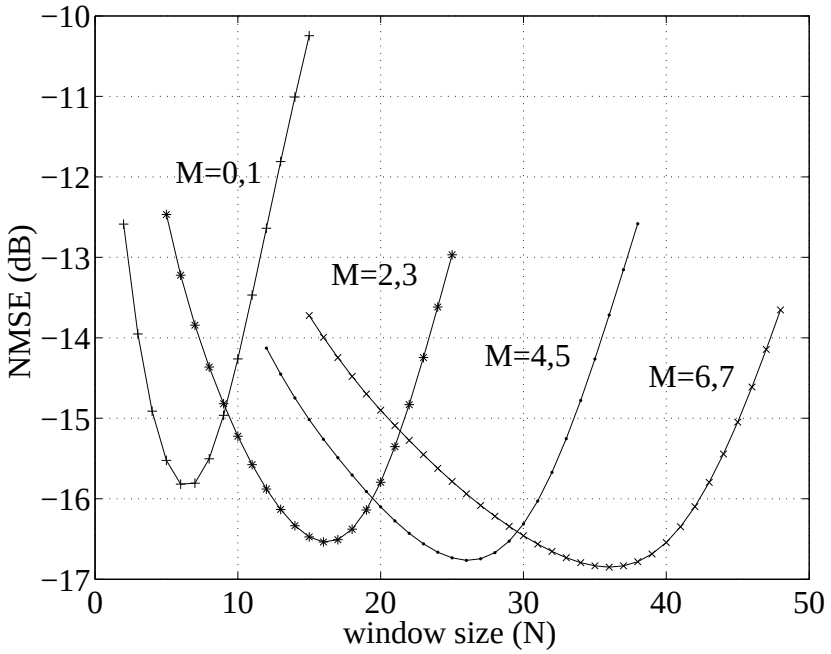


Figure 3.3: NMSE for different window size.

it suggests that the second or third order polynomial is good enough for estimation. Zero-th order polynomial may be used for some applications with low complexity.

4. These simulation results also show that the estimation performance is more sensitive to the window size than to the polynomial polynomial order.

Since the estimation performance is very sensitive to the window size, we investigate this issue next.

Optimal Window Size

If the MMSE criterion is used, the optimal window size can be found to minimize the following MSE of estimation

$$\min_{N} E\left\{\left|MSE = r[n_0] - \mathbf{e}_1(\mathbf{D}_M^T\mathbf{W}\mathbf{D}_M)^{-1}\mathbf{D}_M^T\mathbf{W}\tilde{\mathbf{r}}\right|^2\right\}, \tag{3.23}$$

where the expectation is over the instantaneous correlations and noise. Unfortunately,

there is no close form solution for the problem even if the statistics of the instantaneous correlations and the noise are known. Therefore, we resort to an approximate solution.

In (3.22), we can discard the higher order terms if the window size is well chosen and thus obtain a simplified expression

$$\text{bias}(\hat{\mathbf{a}}_M) \approx N^{M+1} a_{M+1} \bar{\mathbf{N}}_M^{-1} \mathbf{\Psi}_M^{-1} \mathbf{\Phi}_{M+1}, \tag{3.24}$$

where we only consider the case that M is odd for convenience since performance for the even and odd order polynomial-based estimators is same. Substitute (3.21) and (3.24) into (3.11), we obtain the MSE of estimation

$$\varepsilon_p[n_0] \approx \frac{\alpha \tilde{\sigma}^2[n_0]}{N} + \left(\beta a_{M+1} N^{M+1}\right)^2, \tag{3.25}$$

where the constants $\alpha = \mathbf{e}_1 \bar{\mathbf{N}}_M^{-1} \mathbf{\Psi}_M^{-1} \bar{\mathbf{\Psi}}_M \mathbf{\Psi}_M^{-1} \bar{\mathbf{N}}_M^{-1} \mathbf{e}_1^T$ and $\beta = \mathbf{e}_1 \bar{\mathbf{N}}_M^{-1} \mathbf{\Psi}_M^{-1} \mathbf{\Phi}_{M+1}$.Both depend on the polynomial order M and the weight function $W(\cdot)$. The optimal window size can be derived by minimizing the $\varepsilon_p[n_0]$ in (3.25), which is

$$\begin{aligned} N_{opt}^{(p)}[n_0] &= \arg\min_N \varepsilon_p[n_0] \\ &= \left(\frac{\alpha \tilde{\sigma}^2[n_0]}{(2M+2)\beta^2 a_{M+1}^2} \right)^{\frac{1}{2M+3}}. \end{aligned} \tag{3.26}$$

If the order M and the weight function W is determined, coefficients α and β can be calculated. For example, for the weight function of $W(t) = 3/4(1-t^2)$ [66] and $M = 3$, the optimal window size is $1.60(\tilde{\sigma}^2[n_0]/a_{M+1})^{1/2M+3}$.

Sometimes, minimizing the MSE at one point cannot guarantee minimizing the MSE at other points within the window. An alternative approach is to minimize the weighted MSE over the whole window, which is defined as

$$\varepsilon_s = \frac{E\left\{(\mathbf{r} - \mathbf{D}_M \hat{\mathbf{a}}_M)^T \mathbf{W} (\mathbf{r} - \mathbf{D}_M \hat{\mathbf{a}}_M)\right\}}{\text{Tr}(\mathbf{W})}, \tag{3.27}$$

where $\text{Tr}(\mathbf{W})$ denotes the trace of matrix $\mathbf{W}$.

Similar to the derivation of $\varepsilon_p[n_0]$, we can obtain

$$\begin{aligned}\varepsilon_s &\approx \frac{a_{M+1}^2 N^{2M+2}(\psi_{2M+2} - \mathbf{\Phi}_{M+1}^T \mathbf{\Psi}_M^{-1} \mathbf{\Phi}_{M+1})}{\psi_0} \\ &\quad + \frac{\tilde{\sigma}[n_0]^2 \mathrm{Tr}(\bar{\mathbf{\Psi}}_M \mathbf{\Psi}_M^{-1})}{N\psi_0},\end{aligned}$$

where we assume the variance of $\tilde{v}[n]$ is constant. The optimal window size to minimize ε_s is

$$N_{opt}^{(s)} = \left(\frac{\mathrm{Tr}(\bar{\mathbf{\Psi}}_M \mathbf{\Psi}_M^{-1})}{(2M+2)(\psi_{2M+2} - \mathbf{\Phi}_{M+1}^T \mathbf{\Psi}_M^{-1} \mathbf{\Phi}_{M+1})} \right)^{\frac{1}{2M+3}} \left(\frac{\tilde{\sigma}^2[n_0]}{a_{M+1}^2} \right)^{\frac{1}{2M+3}}. \tag{3.28}$$

Comparing (3.26) and (3.28), we can see that two types of the optimal window sizes can be uniformly expressed as

$$N_{opt} = C_{M+1} \left(\frac{\tilde{\sigma}^2[n_0]}{a_{M+1}^2} \right)^{\frac{1}{2M+3}}, \tag{3.29}$$

where C_{M+1} is a constant depending on the order and the weight function.

Polynomial Coefficient Estimation

In the previous section, we have investigated the optimal window size for the polynomial-based estimator. In this section, we study how to obtain polynomial coefficients required to calculate the optimal window size.

The parameter a_{M+1} can be found through pilot estimation. The idea is to use $(M+1)$-th or a higher-order polynomials to fit the temporal estimation and obtain the estimation of a_{M+1} and $\tilde{\sigma}^2$. Thus, an approximately optimal window size can be calculated. From the approximately optimal window size, we use an M-th order polynomial fitting to obtain the refined estimation. The problem for pilot estimation is how to select a proper initial window size.

If statistics of the instantaneous correlations are known, the mean-square of the $M+1$-th order polynomial coefficient a_{M+1} can be obtained by [69]

$$E|a_{M+1}|^2 = \frac{(2\pi)^{2(M+1)}}{(M+1)!^2} \int_{-2f_d}^{2f_d} |f|^{2(M+1)} S_r(f) \mathrm{d}f,$$

where f_d is the Doppler frequency and $S_r(f)$ is the Doppler spectrum of $r[n]$. If we estimate in frequency domain, f_d becomes τ_m, the maximum delay spread, and $S_r(f)$

turns to the frequency correlation spectrum. Then, we can calculate the window size for pilot estimation according to (3.29). For Rayleigh fading channel, and $f_d =$ 100 Hz, $\tilde{\sigma}^2 = 0.1$, $M = 3$, the optimal window size can be calculated as $N_{opt} \approx 14$, and $N_{opt} \approx 17$ for $M = 4$.

However, the statistics of the channels or instantaneous correlations of the received signals are usually not known and need to be estimated. Thus, the global optimal window size based on statistics may not be available. Here, we introduce another approach that does not require the statistics of the channels or instantaneous correlations of the received signals. The basic idea is to find the window size that minimizes some cost function. From [65], the normalized estimation error is defined as

$$\hat{\sigma}^2 = \frac{\sum_{i=n_0-N}^{n_0+N} [r(i) - \hat{r}(i)]^2 W(\frac{i-n_0}{N})}{\mathrm{Tr}(\mathbf{W} - \mathbf{W}\mathbf{D}_M \mathbf{G}_M^{-1} \mathbf{D}_M^T \mathbf{W})}. \tag{3.30}$$

Its expectation is

$$E[\hat{\sigma}^2] = \tilde{\sigma}^2[n_0] + b_q^2, \tag{3.31}$$

where

$$b_q^2 = \kappa^{-1} \boldsymbol{\xi}_M^T \{\mathbf{W} - \mathbf{W}\mathbf{D}_M \mathbf{G}_M^{-1} \mathbf{D}_M^T \mathbf{W}\} \boldsymbol{\xi}_M,$$

and

$$\kappa = \mathrm{Tr}(\mathbf{W} - \mathbf{W}\mathbf{D}_M \mathbf{G}_M^{-1} \mathbf{D}_M^T \mathbf{W}).$$

The second term in (3.31) can be approximated as

$$\begin{aligned} b_q^2 &\approx \kappa^{-1} \{g_{2q+2} - \mathbf{g}_{q+1}^T \mathbf{G}_M^{-1} \mathbf{g}_{q+1}\} a_{q+1}^2 \\ &\approx N^{2q+2} a_{q+1}^2 \psi_0^{-1} (\psi_{2q+2} - \boldsymbol{\Phi}_{q+1}^T \boldsymbol{\Psi}_M^{-1} \boldsymbol{\Phi}_{q+1}) \\ &= \tilde{C}_q N^{2q+2} a_{q+1}^2, \end{aligned} \tag{3.32}$$

where $\mathbf{g}_{q+1} = (g_{q+1}, \cdots, g_{2q+1})^T$, and $\tilde{C}_q = \psi_0^{-1}(\psi_{2q+2} - \boldsymbol{\Phi}_{q+1}^T \boldsymbol{\Psi}_M^{-1} \boldsymbol{\Phi}_{q+1})$.

To find the window size for pilot estimation, we define a cost function,

$$\varepsilon_R = (1 + \lambda/N)\hat{\sigma}^2,$$

where λ is a constant. Note that ε_R is a random variable. Its expectation is,

$$E(\varepsilon_R) \approx \tilde{\sigma}^2[n_0] + b_q^2 + \frac{\lambda\tilde{\sigma}^2[n_0]}{N},$$

where the lower-order term $\frac{\lambda}{N}b_q^2$ with respect to N is discarded. Minimizing $E(\varepsilon_R)$ with respect to N results in the window size

$$N_p = \left(\frac{\lambda\tilde{\sigma}^2[n_0]}{(2q+2)\tilde{C}_q a_{q+1}^2}\right)^{\frac{1}{2q+3}}. \tag{3.33}$$

We can choose the constant λ such that $N_p = N_{opt}$ for a specific order and weight function. And the window size for pilot estimation can be obtained through minimizing ε_R:

$$N_p = \arg\min_N(\varepsilon_R). \tag{3.34}$$

With the initial window size, the polynomial coefficients can be estimated through pilot estimation. The algorithm for polynomial estimation can be summarized as follows:

Algorithm (Two-stage polynomial estimation)

1) Initialization: Set a minimal window size N_{min} and maximal window size N_{max}.

2) Pilot estimation: Fit temporal estimation by a polynomial of order $M+1$ or $M+2$ using the pilot window size N_p in (3.34) and obtain the estimation of $\hat{a}_{M+1}, \hat{a}_{M+2}$ and $\hat{\sigma}^2[n_0]$.

3) Refined estimation: Estimate $r[n_0]$ with the window size obtained by (3.26) or (3.28) using an M-th polynomial.

The disadvantage of the pilot estimation is that performance is bad at the low SNR. if the noise is high, the estimated coefficient a_{M+1} is not accurate, which results in bad window size for refined estimation. It is known that the DFT-based estimator

is robust to high noise because the leakage is small compared to the noise at the low SNR. Therefore, we can use the DFT-based approach for pilot estimation at the low SNR. The detail of DFT-based estimation can be found in [45].

To reduce the computational complexity, we can search window sizes at a few time-frequency points and then average them to obtain a window size that is used for pilot estimation at the rest points. For a two stage estimation, the sensitivity of the refined estimator to the window sizes of pilot estimation is reduced significantly. Another simple method of selecting a window size for the pilot estimation is multiplying a coefficient (usually $1.2 \sim 1.5$) to the average window size of refined estimation. All those methods result in close performance from our simulation results.

3.2.3 Simulation Results

In this section, we demonstrate performance of the polynomial-based estimator through simulation under different environments.

The system model and parameters used in our simulation are the same as those in [22]. The entire bandwidth, 4.096 MHz, is divided into 512 subchannels with subchannel space $\Delta f = 8$ kHz. The symbol duration is $T_s = 1/\Delta f = 125$ μs. An additional 31.25 μs guard interval is used to mitigate intersymbol interference. The 512 subchannels are divided into 32 clusters with a cluster size of 15 and 32 guard subchannels (one guard subchannel between each pair of clusters). Four users transmit data through the channel at the same time and each user accesses eight clusters. Thus, there are 120 subchannels for each user. A (40,20) Reed-Solomon (R-S) code, with each symbol consisting of three QPSK symbols grouped in frequency, is used in the system. Hence, each block forms an R-S codeword. Each time slot contains 10 cluster OFDM blocks with one block used for synchronization and training to suppress error propagation. Interference is assumed to be modulated using the same modulation scheme. At the receiver, the instantaneous correlations and channels

are obtained through the polynomial-based estimator with the adaptive algorithm. After original detection, we use enhanced approach proposed in [45] to improve the estimation. Two types of channel models are used: *typical urban* (TU) and HT.

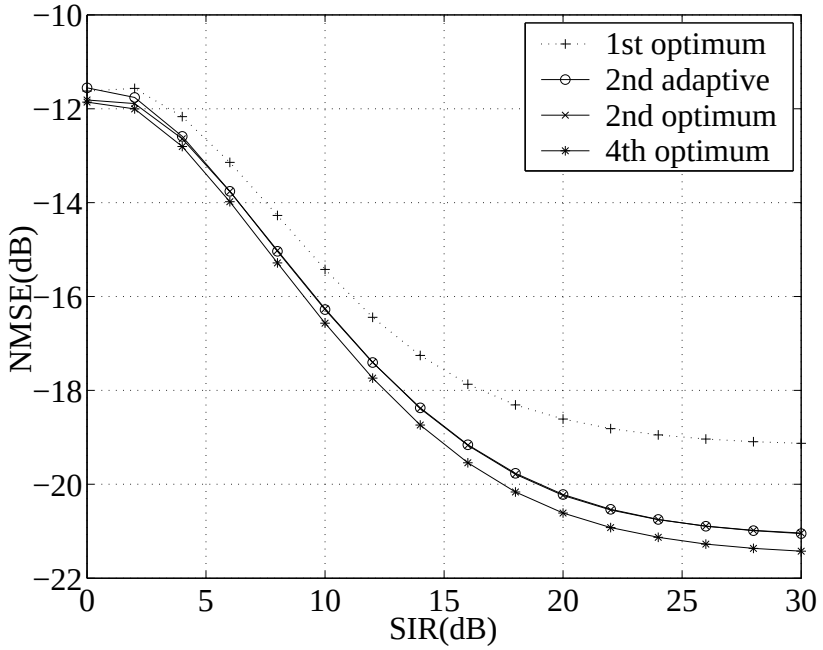
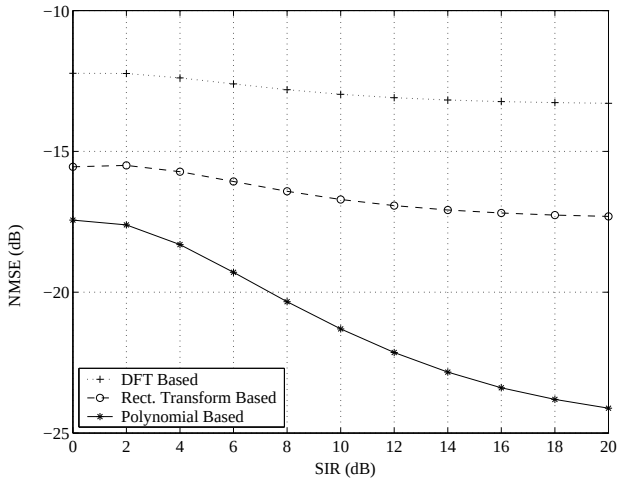


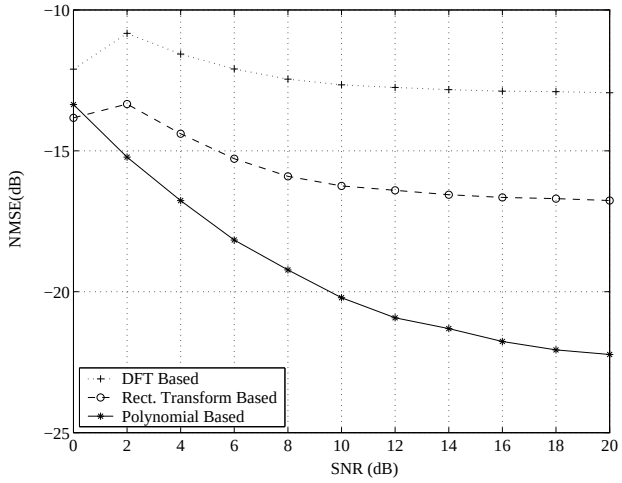
Figure 3.4: NMSE of polynomial-based estimator for channel with $f_d = 100$ Hz at SNR=15 dB.

Figure 3.4 shows the NMSE of the polynomial-based estimator with the optimal window sizes and adaptively adjusted window size for different orders. It can be seen that the performance of the adaptive algorithm is very close to that of the optimal window size for the same order, which shows that the adaptive algorithm performs well even if we do not know the optimal window size. This figure also shows that the second order polynomial approximation is enough for estimation. When the polynomial order increases from two to four, the gain is less than 0.5 dB.

Figures 3.5 and 3.6 compare the NMSE for the DFT-, rectangular- [22], and polynomial-based estimators, under two channel conditions: TU channel with Doppler frequency 40 Hz and HT channel with 100 Hz Doppler frequency. From these figures,



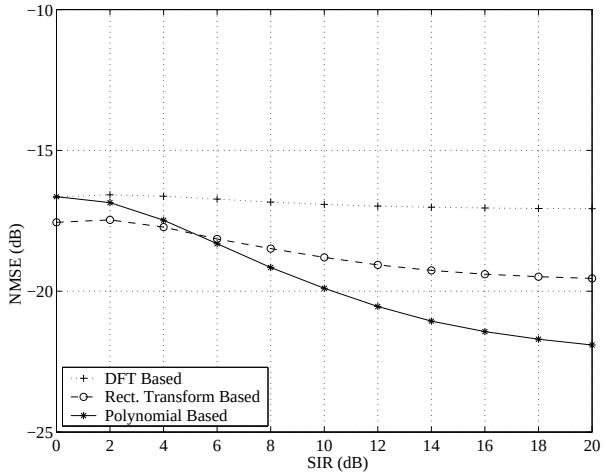
(a)



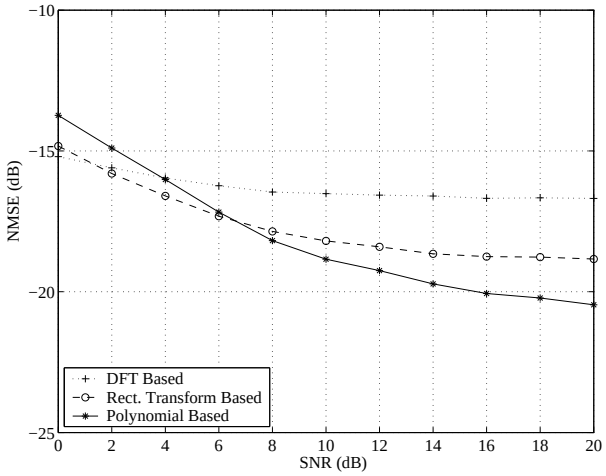
(b)

Figure 3.5: NMSE versus (a) SIR for SNR=15 dB and (b) SNR for SIR=10 dB, for TU channel with $f_d = 40$ Hz.

the polynomial-based estimator has better performance than the other transform-based estimators under the moderate SNR or *signal-to-interference ratio* (SIR). Specifically, from Figure 3.5 for TU channel, there are about a 6.5 dB improvement at 20 dB SIR and 15 dB SNR and a 5.5 dB improvement at 20 dB SNR and 10 dB SIR. With the increase of Doppler frequency or delay spread, performance improvement of the polynomial-based estimator diminishes. For HT channel in Figure 3.6, the performance improvement decreases to 2.5 dB at 20 dB SIR and 15 dB SNR, and



(a)



(b)

Figure 3.6: NMSE versus (a) SIR for SNR=15 dB and (b) SNR for SIR=10 dB, for HT channel with $f_d = 100$ Hz.

1.5 dB at 20 dB SNR and 10 dB SIR, respectively. Because of the leakage effect, the estimation error of the other estimators does not decrease even if the noise and interference are very small. Whereas the polynomial-based estimator has no leakage effect when the window size is adaptively adjusted. Therefore, the performance improvement increases with SINR. However, at lower SINR as in Figure 3.6, there is about a 1 dB performance loss for HT channel. This is because the effect of the interference or noise overwhelms the leakage effect.

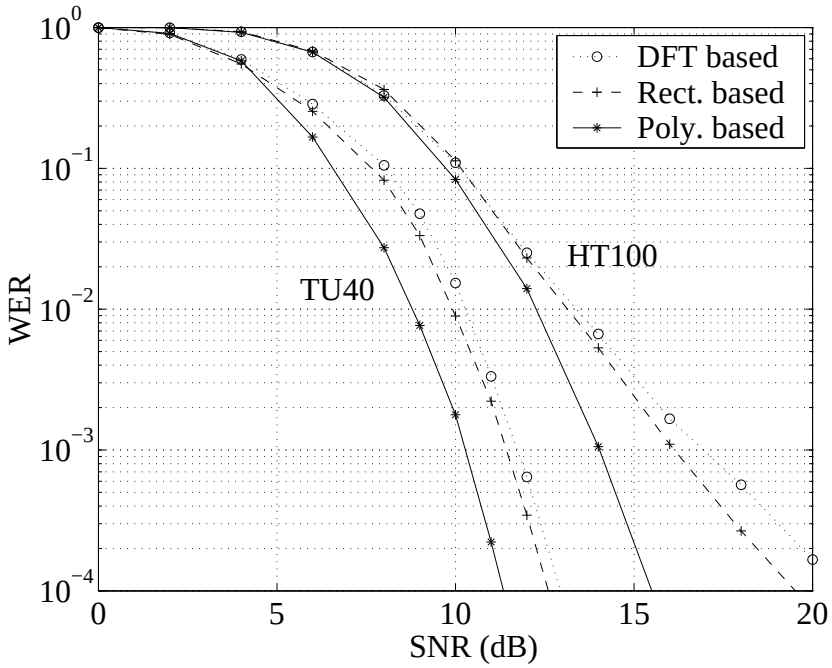


Figure 3.7: WER versus SNR for different channels with SIR=10 dB.

Figures 3.7 and 3.8 show the WER versus SIR and SNR for the DFT-, rectangular-, and polynomial-based estimators, under different channel conditions. Figure 3.8 shows WER versus SIR at 15 dB SNR. Compared to the other estimators, the required SIRs for a 10^{-4} WER for the polynomial-based estimator have about a 1.5 dB and a 2 dB improvement for HT and TU channels, respectively. Figure 3.7 shows the WER versus SNR at 10 dB SIR. For a 10^{-4} WER, the required SNRs for the polynomial-based estimator have about a 2 dB and a 4 dB improvement for TU and HT channels, respectively.

3.3 Clustered OFDM Based Anti-Jamming Modulation

Many military communications require high-data-rate transmission over wireless channels. However, the dispersive fading of wireless channels causes severe ISI and degrades the system performance. OFDM is an effective technique of mitigating ISI.

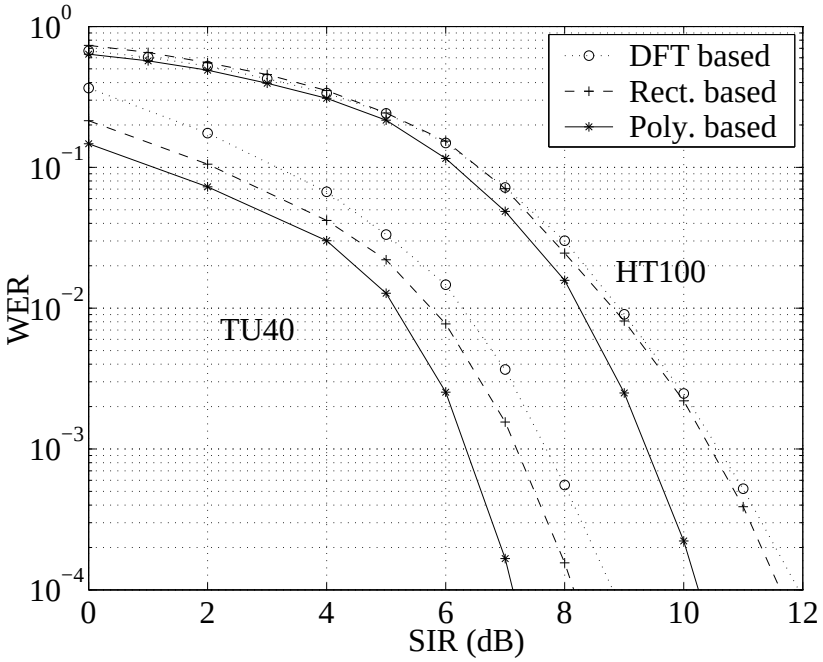


Figure 3.8: WER versus SIR for different channels with SNR=15 dB.

For military communications, jamming exists. Unlike co-channel interference, jamming is intentional interference, which is usually modelled as a Gaussian noise with high power. Here, we consider two types of jamming: *broadband jamming* (BJ) and partial band jamming.

The broadband jamming has the total power J uniformly spread over the whole spread bandwidth W_{ss}. The broadband jamming does not exploit any knowledge of the anti-jamming system except its bandwidth. The performance of anti-jamming system subject to the broadband jamming provides the baseline performance for all of the anti-jamming systems. An effective anti-jamming system should give performance close to or better than the baseline performance, regardless of the type of the jamming [58].

The partial band jamming occurs when the jammer injects the total power into only part of the total bandwidth W_{ss}. Therefore, partial jammer can exploit the power more efficiently provided the prior knowledge about the anti-jamming systems.

The combination of OFDM and spread spectrum, so-called *multicarrier spread spectrum* (MC-SS), can provide strong anti-jamming and ISI suppression capabilities [59]. It was demonstrated in [60] and [61] that MC-SS outperforms single-carrier *direct sequence spread spectrum* (DS-SS).

Recently, we have investigated clustered OFDM for military communications [62]. In this section, we address anti-jamming modulation based on clustered OFDM to provide both anti-jamming and ISI suppression capabilities. We first describe the proposed system and then its performance.

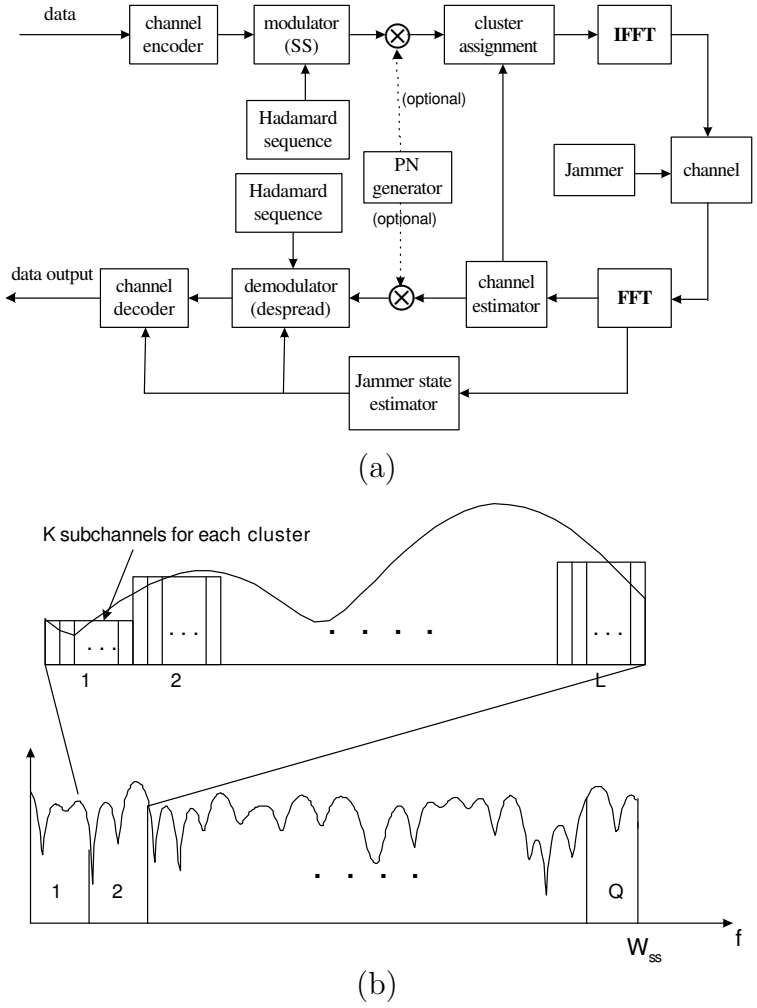


Figure 3.9: Block diagram and cluster assignment for clustered OFDM based spread spectrum systems.

3.3.1 Clustered OFDM Spread Spectrum System

A clustered OFDM SS system for dispersive channels is illustrated in Figure 3.9. In the clustered OFDM SS system, the whole frequency band W_{ss} is uniformly partitioned into Q disjoint subbands with bandwidth $B = W_{ss}/Q$. Each subband includes L OFDM clusters with each one containing K subchannels. Consequently, each subband consists of KL subchannels. Then, the subchannel space is $\Delta f = \frac{W_{ss}}{QKL}$, and the bandwidth of each cluster is $W_c = K\Delta f = \frac{W_{ss}}{QL}$. The cluster bandwidth is carefully selected so that the channel frequency response within each cluster is with a small variation; however, the bandwidth of each subband should be larger than the channel's coherent bandwidth so that the channel frequency responses for different subbands are independent.

Assume that partial channel information is available at the transmitter so that the transmitter is able to choose the cluster with the best quality within each subband. Then, selective diversity can be achieved for combating the fading effect. The clusters assigned to each user *randomly* hop among Q subbands; then, clusters belonging to the same user experience independent fading.

The data stream is first interleaved and encoded with error-correction codes. The encoded data are then spread by a short Hadamard sequence. For security, the spread data symbol may also be scrambled by a pseudo-random noise sequence. Each spread data symbol is modulated using BPSK and then transmitted through one cluster.

Assume a sufficient cyclic prefix is added to suppress the ISI caused by dispersive fading. Then, after DFT and despreading, the received signal for the l-th cluster and n-th block can be expressed as

$$r[n,l] = H[n,l]x[n,l] + J[n,l] + w[n,l], \tag{3.35}$$

where $J[n,l]$ is the jamming signal, $w[n,l]$ is AWGN that is assumed to have zero-mean and variance N_0, $x[n,l]$ is the BPSK modulated signal that is assumed to

have unit variance, and $H[n,l]$ is the channel frequency response that is assumed to be complex Gaussian with zero-mean and unit-variance. $H[n,l]$ is independent for different subbands but has the same statistics.

The jamming signal is assumed to be partial band noise jamming with fraction ρ, $0 < \rho \leq 1$. For convenience of analysis, we assume that the jamming signal occupies multiples of cluster bandwidth. Then, if the total jamming power is P_J Watts, then the received jamming signal $J[n,l]$ has variance $\frac{P_J \triangle f}{2\rho W_{ss}}$ for the cluster with jamming and zero for the cluster without jamming. In the anti-jamming performance analysis, we assume the background noise is negligible compared to the jamming signal if not mentioned.

As indicated before, channel state information, $H[n,l]$ is required at the transmitter to achieve diversity and at the receiver for coherent demodulation and soft-decision decoding. Several channel estimation approaches [22] can be used to estimate channel state information accurately. Therefore, we assume that perfect channel information is available at the transmitter and the receiver when analyzing anti-jamming performance of clustered OFDM SS systems.

3.3.2 Uncoded System Performance

In this section, we investigate performance of clustered OFDM SS systems without coding. We start with systems with broadband jamming, that is, the case of $\rho = 1$.

3.3.2.1 Broadband Jamming Systems

For broadband jamming, the power spectral density of the jamming signal is P_J/W_{ss} and the received jamming power is $N_J/2$ for BPSK modulation, where

$$N_J = \frac{P_J \triangle f}{W_{ss}}.$$

Let $\alpha_l = |H[n,l]|^2$ and

$$\alpha = \max\{\alpha_1, \cdots, \alpha_L\},$$

where we have ignored the time index, n, since it does not affect the analysis. For a given channel gain, α, the *bit-error-rate* (BER) is determined by

$$Pr(e|\alpha) = \mathrm{Q}(\sqrt{\frac{2E_b\alpha}{N_J}}),$$

where E_b is the signal energy per bit and $\mathrm{Q}(x) = \frac{1}{\sqrt{2\pi}}\int_x^{+\infty} e^{-t^2/2}\mathrm{d}t$. Since α is random, the BER for BPSK modulation can be expressed as

$$\begin{aligned} P_b(E_b/N_J) &= \int_0^{\infty} P_\alpha(e|x)\mathrm{d}F_\alpha(x) \\ &= \int_0^{\infty} \frac{1}{2}\sqrt{\frac{E_b}{\pi N_J x}}\exp\left(-\frac{E_b x}{N_J}\right)F_\alpha(x)\mathrm{d}x \end{aligned} \tag{3.36}$$

where $F_\alpha(x)$ is the *cumulative distribution function* (cdf) of α. Since $\alpha = \max\{\alpha_1, \cdots, \alpha_L\}$, the cdf of α can be derived from the joint distribution of $(\alpha_1, \cdots, \alpha_L)$, which is

$$F_\alpha(x) = \int_0^x \int_0^x \cdots \int_0^x f(\alpha_1, \alpha_2, \cdots, \alpha_L)\mathrm{d}\alpha_1\mathrm{d}\alpha_2\cdots\mathrm{d}\alpha_L, \tag{3.37}$$

where $f(\alpha_1, \alpha_2, \cdots, \alpha_L)$ is the joint distribution of $(\alpha_1, \cdots, \alpha_L)$. Let $\mathbf{R}_f$ be the frequency domain channel correlation matrix, defined as

$$\mathbf{R}_f = E\left\{\begin{pmatrix} H[n,1] \\ \vdots \\ H[n,L] \end{pmatrix}(H^*[n,1], \cdots, H^*[n,L])\right\}.$$

For a Rayleigh fading channel, $f(\alpha_1, \alpha_2, \cdots, \alpha_L)$ is a multivariate exponential distribution function that is determined by $\mathbf{R}_f$. The bivariate and trivariate exponential distributions can be obtained from bivariate and trivariate Rayleigh distribution [71]. Unfortunately, there is no close-form expression for arbitrary L and correlation matrices. It is known that there exists a close-form expression for multivariate exponential distribution if the correlation matrix $\mathbf{R}_f$ is an exponential matrix [72], i.e.,

$$(\mathbf{R}_f)_{ij} = \begin{cases} 1, & i = j \\ r_f^{j-i}, & i < j \\ (r_f^*)^{i-j}, & i > j \end{cases}.$$

And the distribution function $f(\alpha_1, \alpha_2, \cdots, \alpha_L)$ can be written as [72]

$$\begin{aligned} f(\quad & \alpha_1, \alpha_2, \cdots, \alpha_L) = \\ & \frac{1}{(1-r^2)^{L-1}} \exp\left(-\frac{1}{1-r^2}\left[\alpha_1 + \alpha_L + (1+r^2)\sum_{i=2}^{L-1} \alpha_i \right] \right) \\ & \times \prod_{i=1}^{L-1} I_0\left(\frac{2r}{1-r^2} (\alpha_i \alpha_{i+1})^{\frac{1}{2}} \right), \end{aligned} \tag{3.38}$$

where $I_0(\cdot)$ is the zeroth-order modified Bessel function of the first kind and $r = |r_f|$ is the correlation coefficient between the adjacent channels. Here, we use the exponential correlation matrix to approximate the practical correlation matrix. Since the distribution is only determined by the absolute value of the correlation coefficient r, we consider only the absolute value in the approximation. Equating $\ln\{|(\mathbf{R}_f)_{1j}|\}$ with $(j-1)\ln r,\ j = 2, \cdots, L-1$ and applying the least-square method, the approximate correlation coefficient between adjacent channels can be obtained as

$$r = \exp\left(\sum_{j=2}^{L} \frac{(j-1)\ln(|(\mathbf{R}_f)_{1j}|)}{L(L-1)(2L-1)/6)} \right). \tag{3.39}$$

To calculate the cdf of the diversity output, we extend the Bessel function into the power series and substitute into (3.38):

$$\begin{aligned} f(\quad & \alpha_1, \cdots, \alpha_L) = \frac{\exp\left(-\frac{1}{1-r^2}\left[\alpha_1 + \alpha_L + (1+r^2)\sum_{i=2}^{L-1} \alpha_i \right] \right)}{(1-r^2)^{L-1}} \\ & \times \sum_{i_1, i_2, \cdots, i_{L-1}=0}^{\infty} \left(\frac{r}{1-r^2} \right)^{2\sum_{j=1}^{L-1} i_j} \\ & \times \frac{\alpha_1^{i_1} \alpha_2^{i_1+i_2} \cdots \alpha_{L-1}^{i_{L-2}+i_{L-1}} \alpha_L^{i_{L-1}}}{\prod_{j=1}^{L-1} (i_j!)^2}. \end{aligned} \tag{3.40}$$

Substituting (3.40) into (3.37) leads to

$$\begin{aligned} F_\alpha(x) \quad & = (1-r^2) \sum_{i_1, i_2, \cdots, i_{L-1}=0}^{\infty} \frac{\prod_{j=1}^{L-1} \binom{i_j + i_{j+1}}{i_j} r^{2\sum_{j=1}^{L-1} i_j}}{(1+r^2)^{i_1 + i_{L-1} + L - 2 + 2\sum_{j=2}^{L-2} i_j}} \\ & \times \gamma(i_1 + 1, \frac{1}{1-r^2} x) \gamma(i_{L-1} + 1, \frac{1}{1-r^2} x) \\ & \times \prod_{j=1}^{L-2} \gamma(i_j + i_{j+1} + 1, \frac{1+r^2}{1-r^2} x), \end{aligned} \tag{3.41}$$

where $\gamma(\cdot)$ is the incomlete Gamma function, which is defined as

$$\gamma(a,x) = \frac{1}{\Gamma(a)} \int_0^x \exp(-t) t^{a-1} \mathrm{d}t,$$

and $\Gamma(\cdot)$ is the gamma function. Then, the BER for an uncoded clustered OFDM SS system can be calculated numerically through (3.36). From the comparison between the numerical and simulation results in the following sections, the approximate performance is very close to the exact one for all *signal-to-jamming ratio* (SJR) regions.

In some cases, the above analytical expressions are too cumbersome to be used in practical system design. And people are more interested in the asymptotic performance at the high SJR. Then, using the method similar to the one in [73], we can obtain a much simpler asymptotic expression.

When SJR tends to infinity, the BER can be expressed by

$$P_{asm}(E_b/N_J) = \lim_{E_b/N_J \to \infty} P_b(E_b/N_J) = \left(G_c \frac{E_b}{N_J} \right)^{-G_d}, \tag{3.42}$$

where G_c is called coding gain and G_d is referred to as diversity gain or diversity order. Expand the cdf of channel output gain α into power series:

$$F_\alpha(x) = \sum_{i=d}^{\infty} a_i x^i.$$

We have

$$G_c = \left(\frac{2\sqrt{\pi}}{\Gamma(d+\frac{1}{2}) a_d} \right)^{\frac{1}{G_d}}, \; G_d = d,$$

and

$$P_{asm}(E_b/N_J) = \left(\left(\frac{2\sqrt{\pi}}{\Gamma(d+\frac{1}{2}) a_d} \right)^{\frac{1}{d}} \frac{E_b}{N_J} \right)^{-d}. \tag{3.43}$$

For our case, from the cdf expression (3.37), it can be easily derived that

$$d = L, \; a_d = \frac{1}{\det(\mathbf{R}_f)}.$$

Then, the asymptotic BER expression is given by

$$P_{asm}(E_b/N_J) = \left(\left(\frac{2\sqrt{\pi}\det(\mathbf{R}_f)}{\Gamma(L+\frac{1}{2})} \right)^{\frac{1}{L}} \frac{E_b}{N_J} \right)^{-L}. \tag{3.44}$$

Specifically, if $\mathbf{R}_f$ is an exponential correlation matrix, $|\mathbf{R}_f| = (1-r^2)^{L-1}$ and the asymptotic BER is

$$P_{asm}(E_b/N_J) = \left(\left(\frac{2\sqrt{\pi}(1-r^2)^{L-1}}{\Gamma(L+\frac{1}{2})} \right)^{\frac{1}{L}} \frac{E_b}{N_J} \right)^{-L}. \tag{3.45}$$

From the above results, the diversity order is independent of the correlation coefficients and is only determined by the number of diversity branches L. The correlation coefficients determine the coding gain G_c. From (3.45), higher correlation results in lower coding gain.

3.3.2.2 Partial Band Jamming Systems

For partial band jamming systems with a jamming fraction ρ, the BER can be similarly derived as

$$P_b(\rho,\ E_b/N_J) = \rho P_b\left(\rho E_b/N_J\right) = \rho \int_0^\infty \frac{1}{2}\sqrt{\frac{E_b\rho}{\pi N_J x}} \exp\left(-\frac{\rho E_b x}{N_J}\right) F_\alpha(x)\mathrm{d}x \tag{3.46}$$

For convenience, we make the following transforms

$$s = \ln(E_b/N_J), P_{\log}(s) = \ln\left(P_b(e^s)\right),$$

where P_b is the BER for broadband jamming. Then, for partial band jamming, the logarithmic BER can be expressed as

$$\begin{aligned} \ln P_b(\rho, \frac{E_b}{N_J}) &= \ln\left(\rho P_b(\rho\frac{E_b}{N_J})\right) \\ &= P_{\log}(s + \ln\rho) + \ln\rho \end{aligned} \tag{3.47}$$

The *worst case jamming* (WCJ) fraction ρ^* can be obtained by maximizing (3.47) with respect to ρ:

$$\rho^* = \max_\rho \ln P_b(\rho, \frac{E_b}{N_J}).$$

Taking the derivative of (3.47) with respect to ρ and making it equal to zero lead to the worst-case jamming fraction ρ^*:

$$\rho^* = \begin{cases} \dfrac{C_L}{E_b/N_J}, & E_b/N_J \geq C_L, \\ 1, & E_b/N_J < C_L, \end{cases} \tag{3.48}$$

where C_L is the solution of the following equation:

$$P'_{\log}(\ln C_L) = -1,$$

where $P'_{\log}(\cdot)$ is the derivative of $P_{\log}(\cdot)$.

For flat Rayleigh fading channels or $L = 1$, the BER is [58]

$$P_b = \frac{1}{2} - \frac{1}{2\sqrt{\frac{N_J}{E_b} + 1}}.$$

It can be calculated that $P'_{\log}(\cdot) > -1$; therefore, the worst-case jamming fraction turns out to be $\rho^* = 1$. For frequency selective fading radio channels and $L > 1$, from the asymptotic expression (3.44), $P'_{\log}(\cdot) = -L < -1$ as $E_b/N_J \to \infty$ and $P'_{\log}(\cdot) = 0 > -1$ as $E_b/N_J \to 0$; then, (3.49) always has a solution. The optimal jamming fraction is less than 1 at the high SJR region. The corresponding worst-case BER is

$$P_w = \begin{cases} \dfrac{C_L}{E_b/N_J} P_b(C_L), & E_b/N_J \geq C_L, \\ P_b(E_b/N_J), & E_b/N_J < C_L. \end{cases} \tag{3.49}$$

Note that the BER for the flat Rayleigh fading channels at large SJR can be approximated as

$$P_b \approx \frac{1}{4E_b/N_J}$$

This shows that the BER of the uncoded systems subject to worst-case jamming decreases with SJR in the same way as that of flat Rayleigh fading channels.

3.3.2.3 Numerical and Simulation Results for Uncoded Systems

In this and the following sections, we use the channel with an exponential delay profile as an example to show the effect of worst-case jamming. The frequency correlation function for an exponential delay profile can be expressed as [22]

$$r_{\exp}(\triangle f) = \frac{\exp(\jmath 2\pi \triangle f \bar{\tau})}{\jmath 2\pi \triangle f \bar{\tau} + 1}, \tag{3.50}$$

where $\bar{\tau}$ is the delay spread. Then, the correlation coefficient between cluster i and $i + l$ is

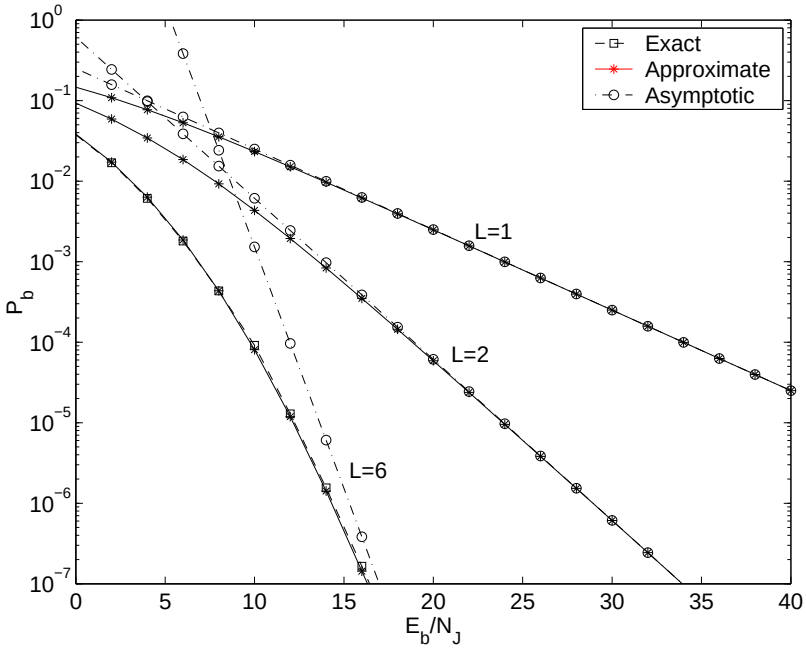


Figure 3.10: Approximate performance for broadband jamming.

$$r_l = \frac{\exp(\jmath 2\pi l W_c \bar{\tau})}{\jmath 2\pi l W_c \bar{\tau} + 1}. \tag{3.51}$$

For $\bar{\tau} W_c = 0.2$, the uncoded system performance is shown in Figures 3.10 and 3.11.

Figure 3.10 gives the exact (simulated), asymptotic, and approximate BER for the broad-band jamming. Compared to the exact result, the approximate BERs are very accurate. Therefore, we can use the approximation in the following sections without

a considerable effect. We can also see that the asymptotic BER curve fits the exact one well at the high SJR region. Then, the simple asymptotic expression can be used when only the performance at high SJR is concerned.

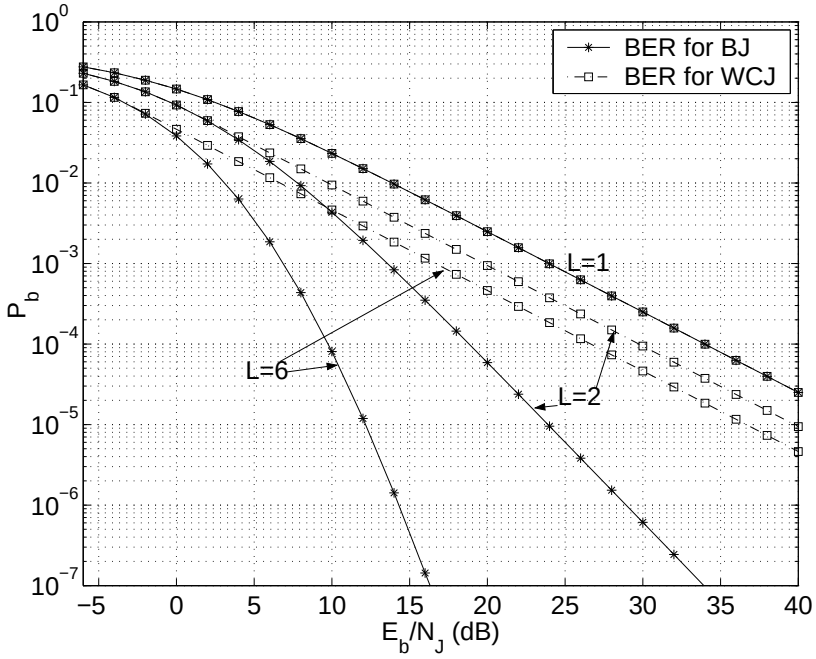


Figure 3.11: Uncoded performance for the worst case jamming and broadband jamming.

Figure 3.11 gives the numerical and simulated BER performance under broad-band jamming and worst-case jamming for $L = 1, 2, 6$. For $L = 1, 2$, the exact BER can be obtained through numerical calculation. For $L = 6$, we use the exponential correlation matrix to approximate the exact one to obtain the BER curve. From the figure, the performance is improved with the increase of L. For broad-band jamming, the required SJR at a 10^{-4} BER is improved by 15 dB and 28 dB, respectively, when L increases from 1 to 2 and 1 to 6. However, it is degraded severely by worst-case jamming. When L increases from 1 to 2 and 6, the performance improvement loses about 11 dB and 21 dB, respectively. Therefore, coding has to be used to recover the loss. In the next section, we analyze the performance of coded systems.

3.3.3 Coded System Performance

Coding is an effective technique to combat jamming and fading. In this section, we first analyze the anti-jamming performance of coded systems with hard-decision decoding.

3.3.3.1 Hard-Decision Decoding

For hard-decision decoding, the channel can be regarded as binary, symmetric, and with a cutoff rate [58]:

$$R_0 = 1 - \log_2(1 + D) \text{ bits/symbol}, \tag{3.52}$$

where parameter

$$D = \sqrt{4P_b(1 - P_b)},$$

with P_b being the uncoded BER. From [58], the performance bound for coded systems with hard-decision decoding is determined by the cutoff rate R_0. Since the cutoff rate depends only on the uncoded BER, the worst-case jamming for uncoded systems is also the worst for the hard-decision decoding system. For a convolutional code with constraint length $K = 7$ and rate $R = 1/2$, the BER is upper bounded by [78]

$$P_b \leq \frac{1}{2}(36D^{10} + 211D^{12} + 1404D^{14} + \cdots). \tag{3.53}$$

The BER of hard-decision decoding systems under the channel with exponential delay profile is shown in Figures 3.12 and Figure 3.13. Figure 3.12 shows the cutoff rate of coded systems with hard-decision decoding. For broad-band jamming, there is a 5 dB improvement when L increases from 1 to 2. With the increase of L, improved performance is obtained. However, the performance loss due to the worst-case jamming still exists at high SJR when $L > 1$. The loss is as large as 4-7 dB when $L = 6$. However, compared to the 24 dB loss of the uncoded system, coding has recovered most of the performance loss.

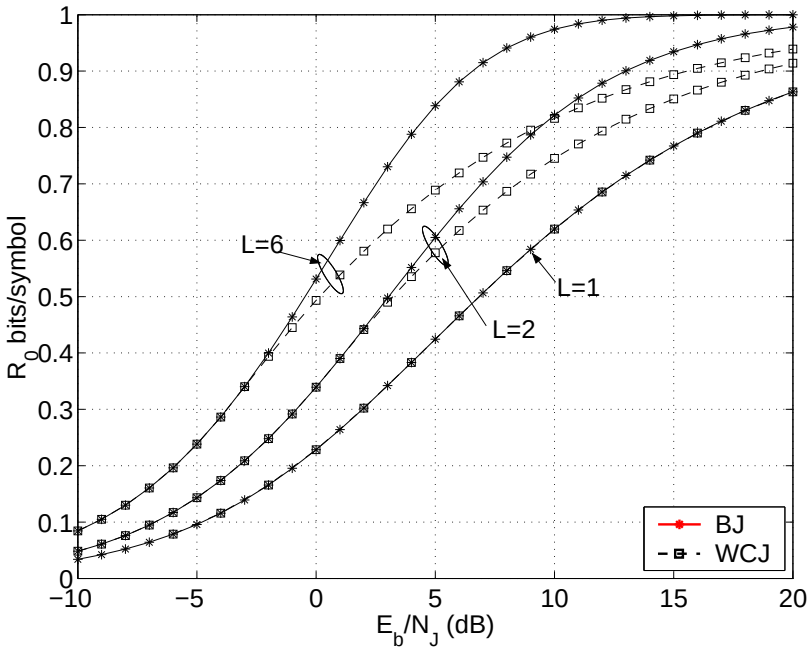


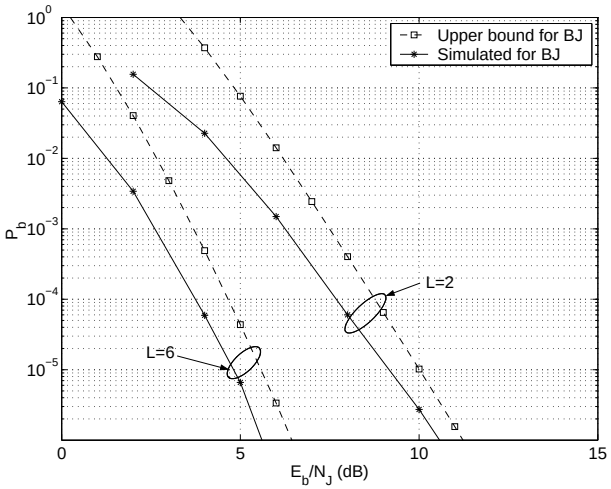
Figure 3.12: Cutoff rate of a coded system with hard-decision decoding for the worst case and broadband jamming.

The BER bounds and simulation results for the coded system with hard-decision decoding are shown in Figure 3.13. From this figure, it is obvious that most of the performance loss because of the worst-case jamming in the uncoded system has been recovered. For example, from $L = 2$ to $L = 6$, the performance loss has been reduced to only $1.5 \sim 3$ dB at a 10^{-5} BER. Compared to the simulation results, the BER bounds have always a 1 dB difference from the exact ones for both broad-band and worst-case jamming.

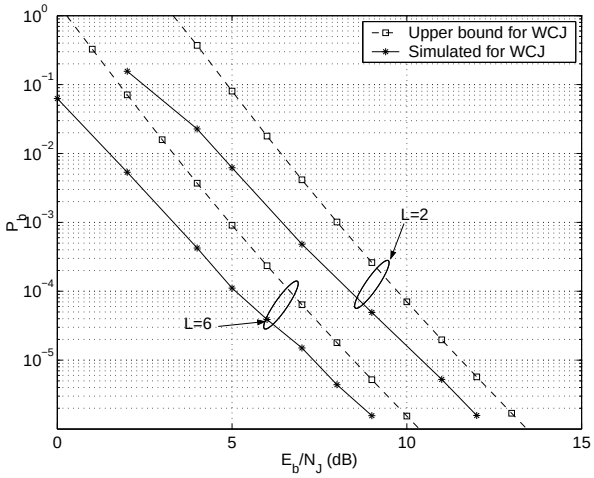
3.3.3.2 Soft-Decision Decoding

Soft-decision decoding can be used to further improve system performance. For systems with soft-decision decoding, we assume that the receiver knows both channel and *jamming state information* (JSI).

For a partial band jamming system, the received signal for one cluster after DFT



(a)



(b)

Figure 3.13: Hard-decision decoded performance for the broadband and worst case jamming.

and despreading can be generally expressed as

$$r = Hx + ZJ, \tag{3.54}$$

where Z is the jamming state, which is equal to one if jammed and zero if unjammed, and H is the channel magnitude after the phase correction. Then, the decision metric for maximum-likelihood detection is

$$m(r, x|H, Z) = c(Z)Hrx, \tag{3.55}$$

where $c(Z)$ is a weight function of jamming state Z. Following the procedures in [58], we can compute the parameter as

$$\begin{aligned} D(\lambda,\rho) &= E\{e^{\lambda[c(Z)rH(\hat{x}-x)]}|x\}|_{\hat{x}\neq x} \\ &= E\left\{\rho e^{-2\lambda c(1)\alpha E_b+\lambda^2 c^2(1)\alpha E_b N_J/\rho}+(1-\rho)e^{-2\lambda c(0)\alpha E_b}\right\}, \end{aligned} \tag{3.56}$$

where $\alpha = |H|^2$ is the channel gain as mentioned above. If we choose $c(1) = 1$ and a sufficiently large $c(0)$ so that the second term in the above equation is negligible, then

$$D(\lambda,\rho) = E\left\{\rho e^{-2\lambda\alpha E_b+\lambda^2\alpha E_b N_J/\rho}\right\}. \tag{3.57}$$

From (3.57), D can be obtained by minimizing $D(\lambda)$ with respect to λ, which is

$$\begin{aligned} D(\rho) &= \min_{\lambda>0} D(\lambda,\rho) \\ &= E\left\{\rho e^{-\frac{\rho\alpha E_b}{N_J}}\right\} \\ &= \rho D_b(\rho E_b/N_J), \end{aligned} \tag{3.58}$$

where D_b represents the D parameter when $\rho = 1$. Then by (3.41) and (3.59), the cutoff rate and BER bound can be calculated from (3.52) and (3.53), respectively. Similar to (3.46), D parameter for partial band jamming can be derived as

$$s = \ln(E_b/N_J),\ D_{\log}(s) = \ln(D_b(e^s)),$$

$$D(\rho) = \rho\int_0^\infty \frac{E_b\rho}{N_J}\exp(-\frac{E_b\rho x}{N_J})F_\alpha(x)\mathrm{d}x. \tag{3.59}$$

Making the transforms

Then, similar to the derivation of worst case jamming fraction for uncoded systems, the worst case jamming fraction can be expressed the same as (3.48), where the constant C_L is the solution of following equation

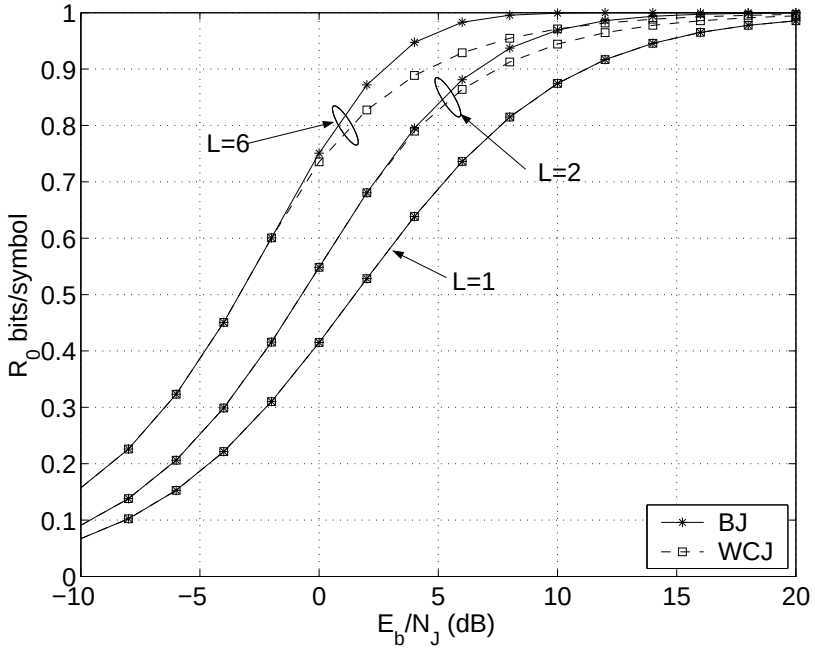


Figure 3.14: Cutoff rate with soft-decision decoding for the worst case and broadband jamming.

For flat fading or $L = 1$, the parameter D can be written in close-form as

$$D(\rho) = \frac{\rho}{\rho E_b/N_J + 1}. \tag{3.60}$$

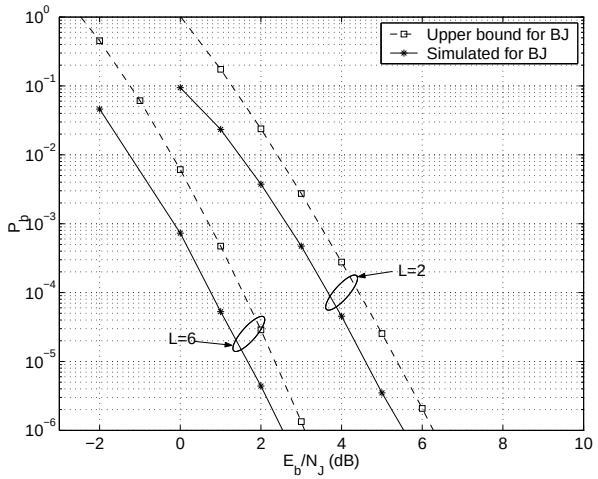
At high SJR, D_ρ is approximated as

$$D(\rho) \approx \frac{1}{E_b/N_J}, \tag{3.61}$$

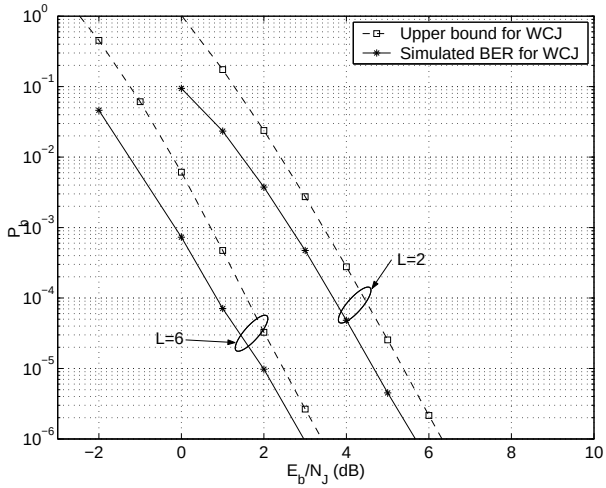
and the worst-case jamming fraction is $\rho^* = 1$. When $L > 1$, we have $\rho < 1$ at high SJR.

$$D'_{\log}(\ln C_L) = -1. \tag{3.62}$$

Figures 3.14 and 3.15 show the numerical and simulation results for coded systems with soft-decision decoding. From the figures, the performance loss due to the worst-case jamming is almost recovered by soft-decision decoding. The jammer is forced to spread its energy over the whole band. It also shows that soft-decision decoding has



(a)



(b)

Figure 3.15: Soft-decision decoded performance for the broadband and worst case jamming.

a much better anti-jamming ability than hard-decision decoding. For example, for the case of $L = 2$, there is about a 5 dB improvement at a 10^{-5} BER. Again, about a 1 dB difference is observed between the exact BER and the upper bound.

3.3.3.3 Jamming State Estimation

From the previous section, to recover performance loss due to jamming, soft-decision decoding should be used, which, however, needs channel and jamming state information. Hence, we investigate jamming state estimation for a coded system with

soft-decision decoding.

Consider the received signal vector for one cluster after despreading:

$$\mathbf{r} = H\mathbf{x} + Z\mathbf{J} + \mathbf{w}, \tag{3.63}$$

where $\mathbf{J} = (J_1, \cdots, J_K)^T$ with $\mathrm{Var}(J_k) = N_J/(2\rho)$, H is the channel magnitude after the phase correction, and Z is the jamming state, $Z \in \{0, 1\}$. For analysis convenience, we assume the whole cluster is used to transmit data. The jamming state detector calculates the log-likelihood of Z by

$$\Lambda(Z) = \sum_{k=1}^{K} \ln \frac{Pr(r_k|Z=1)}{Pr(r_k|Z=0)}. \tag{3.64}$$

Through some manipulations, we have

$$\begin{aligned} \ln \frac{Pr(r_k|Z=1)}{Pr(r_k|Z=0)} &= \frac{r_k^2 + H^2}{N_0(\rho N_0/N_J + 1)} \\ + \ln\left(\cosh \frac{2r_k H}{N_0 + \frac{N_J}{\rho}}\right) &- \ln\left(\cosh \frac{2r_k H}{N_0}\right) \\ &- \frac{1}{2}\ln\left(1 + \frac{N_J}{\rho N_0}\right). \end{aligned} \tag{3.65}$$

Using approximation $\ln(\cosh(|x|)) \approx |x| - C$ for $|x| \geq 1$, the log-likelihood can be simplified as

$$\ln \frac{Pr(r_k|Z=1)}{Pr(r_k|Z=0)} \quad = \frac{(|r_k| - |H|)^2}{N_0(\rho N_0/N_J + 1)} - \frac{1}{2}\ln\left(1 + \frac{N_J}{\rho N_0}\right)$$

Define $\Lambda^* = \sum_{k=1}^{K}(|r_k| - |H|)^2$, then the jamming state can be estimated as follows

$$\begin{cases} Z = 1, & \text{if} \quad \Lambda^* \geqslant \theta \\ Z = 0, & \text{if} \quad \Lambda^* < \theta, \end{cases}$$

where threshold θ is

$$\theta = \frac{KN_0(\rho N_0/N_J + 1)}{2}\ln\left(1 + \frac{N_J}{\rho N_0}\right), \tag{3.66}$$

for the maximum-likelihood detection.

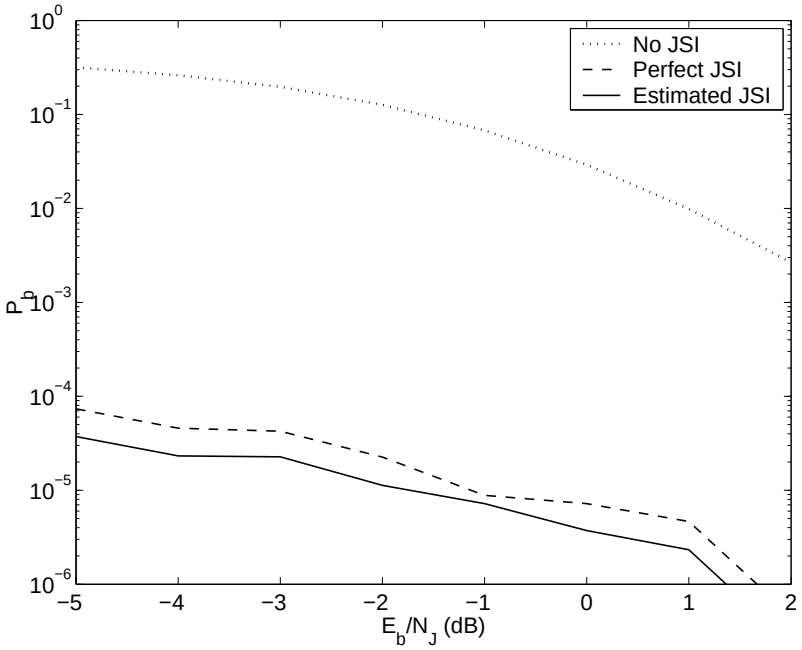


Figure 3.16: Soft-decision decoding performance with jamming fraction $\rho = 0.3$.

Figure 3.16 shows the soft-decision decoding performance with the estimated JSI, where the jamming fraction is $\rho = 0.3$, $K = 16$, and the signal-to-background-noise ratio is set to be 15 dB. For comparison, we also give the performance of soft-decision decoding with perfect JSI and without JSI. As we can see, the receiver with the estimated JSI performs very well. Without the jamming state information, the soft-decision decoding does not provide good performance when the jamming fraction is small. It is noted that the performance with the estimated JSI is even better than that with perfect JSI. This is because the perfect JSI is in the sense of probability and does not reflect the instantaneous jamming signal amplitude while JSI obtained by the jamming state estimator shows the exact amplitude of the jamming signal. Figure 3.17 shows the BER at the worst case jamming cases. From the figure, the BER for the system with the estimated JSI is very close to the one with the exact JSI. Therefore, the simplified jamming state estimator can be used for the soft decision decoding to improve the anti-jamming capability.

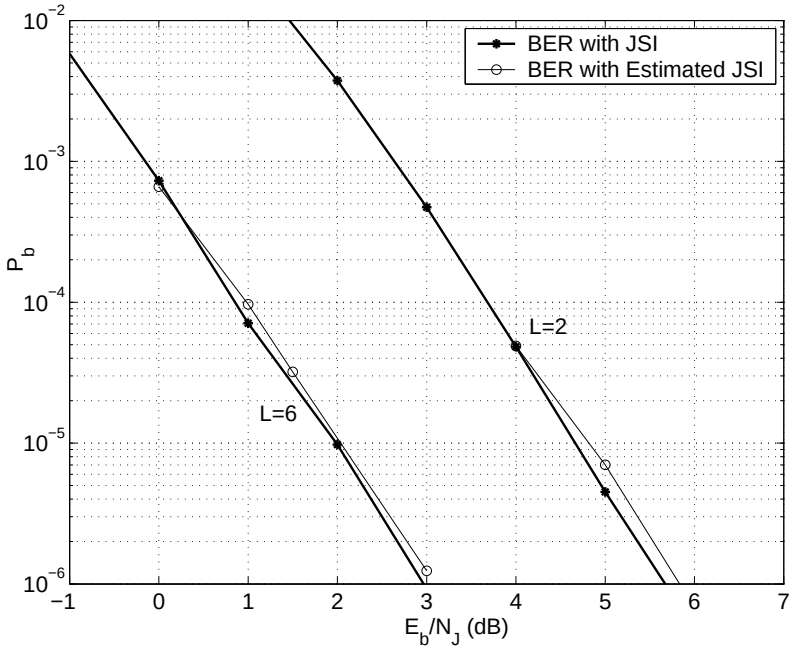


Figure 3.17: Soft-decision decoding performance with estimated JSI.

3.4 Summary

In this chapter, we investigated the co-channel suppression for clustered OFDM and proposed a clustered OFDM based spread spectrum system for military applications. A polynomial-based parameter estimator is proposed to combat severe leakage of the traditional DFT-based estimator in the clustered OFDM with receive antenna arrays for interference suppression. For a clustered OFDM based spread spectrum system, we use an approximate formula to analyze the uncoded system performance. The simulation results show that the approximate BER is very close to the exact one. For coded systems, we analysis their performance bounds and cutoff rates. Our research results show that clustered OFDM spread spectrum system with channel coding can effectively mitigate jamming and fading effect.

3.5 Appendix: Proof of the Stattement in Section 3.2.2.1

To prove the statement, we use the following lemma

Lemma: Let matrix $\mathbf{\Psi}_M^{-1} = \{\rho_{ij}\}_{0\le i,j\le M}$ and let $\mathbf{\Upsilon}_M^1$ be the first row of $\mathbf{\Psi}_M^{-1}$. Then, we have $\rho_{ij} = 0$ if $i+j$ is odd. Furthermore, the first row of $\mathbf{\Psi}_M^{-1}$ and $\mathbf{\Psi}_{M+1}^{-1}$ have the relationship of $\mathbf{\Upsilon}_{M+1}^1 = [\mathbf{\Upsilon}_M^1, 0]$ for even M.

Proof: When $M = 0$, we have $\mathbf{\Psi}_M^{-1} = \psi_0^{-1}$, and

$$\mathbf{\Psi}_{M+1}^{-1} = \begin{pmatrix} \psi_0^{-1} & 0 \\ 0 & \psi_2^{-1} \end{pmatrix}.$$

Assume that the lemma holds for $M = 2k$, we will prove that it holds for $M = 2k+1$ and $M = 2k+2$. Partition $\mathbf{\Psi}_{M+1}$ into block matrices,

$$\mathbf{\Psi}_{M+1} = \begin{pmatrix} \mathbf{\Psi}_M & \mathbf{\Phi}_{M+1} \\ \mathbf{\Phi}_{M+1}^T & \psi_{2M+2} \end{pmatrix}. \tag{3.A.1}$$

Using block matrix inversion, we have

$$\mathbf{\Psi}_{M+1}^{-1} = \begin{pmatrix} \mathbf{\Psi}_M^{-1} + \mathbf{\Psi}_M^{-1}\mathbf{\Phi}_{M+1}S^{-1}\mathbf{\Phi}_{M+1}^T\mathbf{\Psi}_M^{-1} & -\mathbf{\Psi}_M^{-1}\mathbf{\Phi}_{M+1}S^{-1} \\ -S^{-1}\mathbf{\Phi}_{M+1}^T\mathbf{\Psi}_M^{-1} & S^{-1} \end{pmatrix} \tag{3.A.2}$$

where $S = \psi_{2M+2} - \mathbf{\Phi}_{M+1}^T\mathbf{\Psi}_M^{-1}\mathbf{\Phi}_{M+1}$ is a scalar. Note that the vector $\mathbf{\Phi}_{M+1}$ has the form of $(0, *, \cdots, *, 0)$. Therefore, $\mathbf{\Psi}_M^{-1}\mathbf{\Phi}_{M+1}$ will be in the form of $(0, *, 0, *, \cdots, 0, *, 0)$, where $(*)$ denotes any number that is not necessarily zero. Substituting it into (3.A.2) shows the lemma holds for $M = 2k+1$. Similarly, it can be proved for $M = 2k+2$. Thus, the lemma holds for any M.

With the lemma, and the property $\bar{\psi}_{2k+1} = 0$, it can be obtained that the first diagonal elements of matrices $\bar{\mathbf{N}}_M^{-1}\mathbf{\Psi}_M^{-1}\bar{\mathbf{\Psi}}_M\mathbf{\Psi}_M^{-1}\bar{\mathbf{N}}_M^{-1}$ and $\bar{\mathbf{N}}_{M+1}^{-1}\mathbf{\Psi}_{M+1}^{-1}\bar{\mathbf{\Psi}}_{M+1}\mathbf{\Psi}_{M+1}^{-1}\bar{\mathbf{N}}_{M+1}^{-1}$ for even M are the same. Thus, the estimation variance in (3.21) is same for M-th and $(M+1)$-th order polynomials. Also with similar derivation, the estimators using the M-th and $M+1$-th order polynomial have the same bias for even M according to (3.22) and the lemma. According to (3.11), the statement is proved.

CHAPTER IV

MIMO-OFDM FOR WIRELESS LAN

4.1 MIMO-OFDM

The use of multiple antennas at both transmitter and receiver to form a *multiple input multiple output* (MIMO) channel has drawn considerable attention for its potential to improve link reliability and increasing capacity in wireless communications [47]-[34]. In rich scattering propagation environments, for a narrow-band communication system with M_T transmit antennas and M_R receive antennas, it is well known [47], [48] that the channel capacity increases linearly with the minimal number of transmit and receive antennas, $\min(M_T,\ M_R)$.

For broadband communications, OFDM turns a frequency selective channel into a set of parallel flat channels, which significantly reduces the receiver complexity. As a result, the combination of MIMO techniques with OFDM (MIMO-OFDM) is regarded as a promising solution to enhance the data rate of future broadband wireless communication systems. Besides the spatial diversity, broadband channels offer high capacity and frequency diversity resulting from the delay spread [79]. Various schemes have been proposed for MIMO-OFDM systems to exploit spatial-frequency diversity [80], [81]. Recently, MIMO-OFDM has been proposed for the standard of the next generation wireless LAN.

In this chapter, we investigate MIMO-OFDM for wireless LAN. We are exploiting the channel characteristics in spatial and frequency domain to improve performance of channel estimation. For slow varying wireless channel such as indoor channels, partial channel information can be obtained through feedback channel. Then, the

performance of MIMO systems can be improved substantially by exploiting the channel information at the transmitter. In Section 4.2, channel estimation of MIMO-OFDM for wireless LAN is presented. In Section 4.3, linear precoding schemes for MIMO-OFDM with limited channel feedback are discussed.

Here are some notations that will be used in this chapter.

$\mathbf{X}$	$P \times Q$ matrix
$\mathbf{x}_i$	i-th column of $\mathbf{X}$
$r(\mathbf{X})$	rank of $\mathbf{X}$
$\lambda\{\mathbf{X}\}$	eigen spectrum of $\mathbf{X}$
$\mathbf{X}^T$	transpose of $\mathbf{X}$
$\mathbf{X}^*$	complex conjugate of $\mathbf{X}$
$\mathbf{X}^H$	conjugate transpose of $\mathbf{X}$
$\mathbf{X}^{-1}$	inverse of $\mathbf{X}$
$\mathrm{tr}(\mathbf{X})$	trace of $\mathbf{X}$
$\det(\mathbf{X})$	determinant of $\mathbf{X}$
$\otimes$	Kronecker product
$\odot$	Hardamard product
$\mathbf{vec}(\mathbf{X})$	$[\mathbf{x}_1^T, \mathbf{x}_2^T, \cdots, \mathbf{x}_Q^T]^T$
$\mathrm{diag}([x_1, x_2, \cdots, x_Q])$	diagonal matrix with diagonal elements $x_1, x_2, \cdots, x_Q$
$\mathrm{diag}\left([\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_Q]\right)$	block diagonal matrix with diagonal submatrices $\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_Q$
$E\{\cdot\}$	expectation
$\mathbf{I}_P$	$P \times P$ identity matrix
$\mathbf{0}_P$	$P \times P$ zero matrix
$z \sim \mathcal{CN}(0, \sigma^2)$	circularly symmetric complex Gaussian random variable with zero mean and variance σ^2.

4.2 Channel Estimation in Correlated Fading Channels

Most of MIMO transmission schemes require channel state information at the receiver for signal detection. In practice, the channel state information is obtained through channel estimation. Channel estimation for OFDM system with single transmit antenna has been addressed in [38], where the time and frequency correlations of dispersive fading channels are exploited to improve the performance. In MIMO-OFDM systems, channel estimation is more challenging since more parameters have to be estimated. Under the assumption that MIMO channels are *i.i.d.* Rayleigh fading, channel estimators for MIMO-OFDM system have been developed in [55]-[83], where the least square (LS) approach is used and the training or pilot symbols are optimized to minimize the estimation error. However, the MIMO channels are correlated in certain environments, such in indoor mibile systems [56], [57]. Then, channel estimation can achieve better performance by making use of the spatial fading correlation.

In this section, we address channel estimation for spatially correlated broadband MIMO channels. Since statistics of channels are relative stable, we assume that the correlation matrices are known at the transmitter and receiver. With the knowledge of correlation matrices, MMSE channel estimation can be developed. We also investigate the conditions for the optimum training sequences that minimize the estimation error and simplify the estimation. We provide two training sequence design approaches with constant modulus. Both have the same estimation error and result in a simple estimation structure. The optimal window size and polynomial order are studied for the estimator.

4.2.1 System Model for MIMO-OFDM

A broadband MIMO-OFDM system model is shown in Figure 4.1. The data stream $\mathbf{b}(n)$ is first encoded by a space-time or space-frequency encoder. Then, the coded

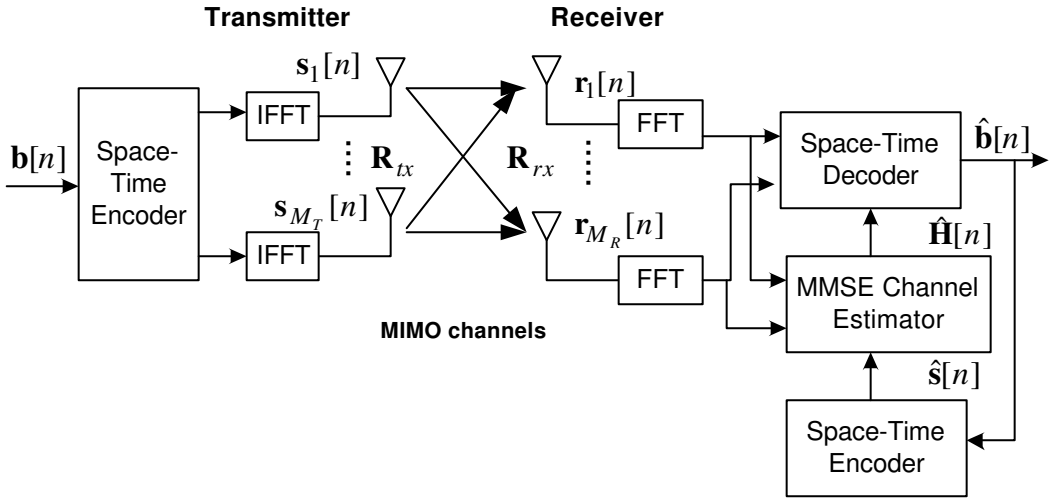


Figure 4.1: Broadband MIMO-OFDM system Model.

data is divided into M_T substreams with each substream forming an OFDM block transmitted through one transmit antenna. At the receiver, the received signals at multiple receive antennas are decoded using channel state information obtained through a training-based or decision-directed estimator. Therefore, channel estimation is critical to achieve the advantages of a MIMO system.

Assume that there are N subchannels for each OFDM block and the length of the cyclic prefix is greater than the delay spread so that there is no ISI. For convenience, we assume a discrete time channel model. The leakage effect caused by the model is negligible for proper cyclic prefix and sampling time [38]. The channel frequency response at n-th block k-th subchannel for a MIMO-OFDM system can be expressed as

$$\mathbf{H}(n,k) = \sum_{l=0}^{L-1} \mathbf{H}_l(n) e^{-\jmath 2\pi lk/N}, \tag{4.1}$$

where $\mathbf{H}_l(n)$ is the l-th tap channel matrix at the n-th block. The received signal at k-th subchannel is given by

$$\mathbf{r}(n,k) = \sqrt{\frac{\rho}{M_T}} \mathbf{H}(n,k)\mathbf{s}(n,k) + \mathbf{w}(n,k), \tag{4.2}$$

where ρ is the SNR,

$$\mathbf{r}(n,k) = [r_0(n,k), r_1(n,k), \cdots, r_{M_R-1}(n,k)]^T$$

is the received signal vector, $\mathbf{H}(n,k)$ is frequency domain channel matrix,

$$\mathbf{s}(n,k) = [s_0(n,k), s_1(n,k), \cdots, s_{M_T-1}(n,k)]^T$$

is the transmitted signal vector with unit variance,

$$\mathbf{w}(n,k) = [w_0(n,k), w_1(n,k), \cdots, w_{M_R-1}(n,k)]^T$$

is AWGN with zero mean and unit variance. Substituting (4.1) into (4.2), and rearranging it, we can obtain

$$\mathbf{r}(n,k) = \sqrt{\frac{\rho}{M_T}}\mathbf{A}(n,k)\mathbf{h}(n) + \mathbf{w}(n,k), \tag{4.3}$$

where

$$\mathbf{h}(n) = [\mathbf{vec}(\mathbf{H}_0)^T, \mathbf{vec}(\mathbf{H}_1)^T, \cdots, \mathbf{vec}(\mathbf{H}_{L-1})^T]^T$$

is the vector of discrete-time channel responses,

$$\mathbf{A}(n,k) = [\mathbf{D}_0(n,k), \mathbf{D}_1(n,k), \cdots, \mathbf{D}_{L-1}(n,k)]$$

with $\mathbf{D}_l(n,k) = e^{-\jmath 2\pi \frac{kl}{N}}\mathbf{s}^T(n,k) \otimes \mathbf{I}_{M_R}$. From (4.3), we can arrange the received signals at all subchannels into vector form:

$$\mathbf{r}(n) = [\mathbf{r}(n,0)^T, \mathbf{r}(n,1)^T, \cdots, \mathbf{r}(n,N-1)^T]^T,$$

and express it uniformly as

$$\mathbf{r}(n) = \sqrt{\frac{\rho}{M_T}}\mathbf{A}(n)\mathbf{h}(n) + \mathbf{w}(n), \tag{4.4}$$

where

$$\mathbf{A}(n) = [\mathbf{A}(n,0)^T, \mathbf{A}(n,1)^T, \cdots, \mathbf{A}(n,N-1)^T]^T$$

and

$$\mathbf{w}(n) = [\mathbf{w}(n,0)^T, \mathbf{w}(n,1)^T, \cdots, \mathbf{w}(n,N-1)^T]^T.$$

4.2.2 Correlated Broadband MIMO Channel Model

We follow the channel model presented in [79] that uses transmit and receive correlation matrices to characterize the statistics of the MIMO channel. Assume that there are M_T transmit antennas and M_R receive antennas, and each channel has L resolvable taps. The $M_R \times M_T$ MIMO channel matrix $\mathbf{H}_l$ of the l-th tap at one instance of time can be represented as

$$\mathbf{H}_l = \sigma_l \left(\sqrt{\frac{K_l}{K_l+1}} \bar{\mathbf{H}}_l + \sqrt{\frac{1}{K_l+1}} \tilde{\mathbf{H}}_l \right), \tag{4.5}$$

where $\bar{\mathbf{H}}_l$ is a fixed matrix that represents the expectation of l-th channel matrix, $\tilde{\mathbf{H}}_l$ is a random matrix with entries being correlated zero-mean, unit variance, complex Gaussian distributed. K_l is the Ricean K-factor, and σ_l^2 is the power of l-th tap, and the total power is normalized to one, $\sum_{l=0}^{L-1} \sigma_l^2 = 1$. Each channel tap can be regarded as a significant scatterer cluster [79]. We assume that coefficients belong to the same tap have the same delay power, and different channel taps are uncorrelated,

$$E\left\{ \mathbf{vec}(\tilde{\mathbf{H}}_l) \left(\mathbf{vec}(\tilde{\mathbf{H}}_{l'}) \right)^H \right\} = \mathbf{0}_{M_R M_T}, \text{ for } l \neq l'. \tag{4.6}$$

Assume that the Ricean component is known, we only need to estimate $\tilde{\mathbf{H}}_l$. Then, we set K_l's to be zero and the channel is purely Rayleigh, $\mathbf{H}_l = \sigma_l \tilde{\mathbf{H}}_l$. For the same tap, the channel coefficients between any antenna pairs are correlated. The spatial correlation can be modelled by decomposing the Rayleigh matrix into

$$\mathbf{vec}(\mathbf{H}_l) = \mathbf{R}_l^{1/2} \mathbf{vec}(\mathbf{H}_{wl}), \tag{4.7}$$

where $\mathbf{H}_{wl}$ is a $M_R \times M_T$ matrix, with $i.i.d$ entries, $\{\mathbf{H}_{wl}\}_{i,j} \sim \mathcal{CN}(0,1)$, $\mathbf{R}_l = \mathbf{R}_l^{1/2} \mathbf{R}_l^{1/2}$ is the correlation matrix of channel vector $\mathbf{vec}(\tilde{\mathbf{H}}_l)$. Although the above correlation matrix can fully capture the correlation between any channel pairs, a simpler model is efficient [84],

$$\mathbf{H}_l = \mathbf{R}_{rl}^{1/2} \mathbf{H}_{wl} \mathbf{R}_{tl}^{1/2}, \tag{4.8}$$

where $\mathbf{R}_{rl}$ and $\mathbf{R}_{tl}$ are the receive and transmit correlation matrices at the l-th tap, respectively. The tap power σ_l^2 is incorporated into the correlation matrix. This model assumes that the receive correlation is same for all transmit antennas and the transmit correlation is same for all receive antennas. The correlation between two distinct channel connecting two different antennas at both transmitter and receiver is equal to the product of the transmit and receive correlations. The transmit and receive correlation matrices of each tap can be determined by power angular spectrum of the scatter cluster [84]. Assume that the *angle of departure* (AOD) and *angle of arrival* (AOA) present Gaussian distribution. Let θ_{tl} be the angle of departure with mean $\bar{\theta}_{tl}$ and variance $\sigma^2_{\theta_{tl}}$, $\bar{\theta}_{rl}$ be the angle of arrival with mean $\bar{\theta}_{rl}$ and variance $\sigma^2_{\theta_{rl}}$. For a *uniform linear array* (ULA), the correlation coefficients between transmit antennas with relative spacing Δ_T and receive antennas with relative spacing Δ_R can be expressed as $\rho(\Delta_T, \bar{\theta}_{tl}, \sigma^2_{\theta_{tl}})$ and $\rho(\Delta_R, \bar{\theta}_{rl}, \sigma^2_{\theta_{rl}})$, respectively. The relative spacing between antennas is defined as $\Delta = d/\lambda$, where d is the absolute antenna spacing and λ is the wavelength. The relation between full correlation matrix $\mathbf{R}_l$ and transmit-receive correlation matrices is

$$\mathbf{R}_l = \mathbf{R}_{rl} \otimes \mathbf{R}_{tl}. \tag{4.9}$$

Then for MIMO-OFDM systems, by (4.6) and (4.9), the correlation matrix of channels $\mathbf{h}(n)$ in (4.4)is given by

$$\begin{aligned} \mathbf{R}_h &= E\{\mathbf{h}(n)\mathbf{h}(n)^H\} \\ &= \text{diag}([\mathbf{R}_{r0} \otimes \mathbf{R}_{t0}, \mathbf{R}_{r1} \otimes \mathbf{R}_{t1}, \cdots, \mathbf{R}_{r(L-1)} \otimes \mathbf{R}_{t(L-1)}]). \end{aligned} \tag{4.10}$$

4.2.3 Basic Channel Parameter Estimation

The spatial correlation in the real MIMO channels is usually changing very slowly compared with the instantaneous channel, therefore, we assume that the spatial correlation is static and known at the transmitter and receiver.

Define

$$\mathbf{h}_f(n) = \left[\mathbf{vec}(\mathbf{H}(n,0))^T, \mathbf{vec}(\mathbf{H}(n,1))^T, \cdots, \mathbf{vec}(\mathbf{H}(n,N-1))\right]^T$$

as the frequency response vector. For MMSE estimation, during training period, the transmitted signal is known and the temporal estimation of $\mathbf{h}_f(n)$ can be obtained by minimizing the MSE

$$\begin{aligned} \text{MSE} &= E\left\{\left|\left|\hat{\mathbf{h}}_f(n) - \mathbf{h}_f(n)\right|\right|^2\right\} \\ &= E\left\{\left(\hat{\mathbf{h}}_f(n) - \mathbf{h}_f(n)\right)^H\left(\hat{\mathbf{h}}_f(n) - \mathbf{h}_f(n)\right)\right\}. \end{aligned} \tag{4.11}$$

Let $\mathbf{F}$ be the $N \times N$ unitary DFT matrix, the relationship between time domain channel responses and frequency domain channel responses is

$$\mathbf{H}_v(n) = \sqrt{N}\left(\mathbf{F} \otimes \mathbf{I}_{M_R M_T}\right)\left[\mathbf{h}^T(n), 0, \cdots, 0\right]^T$$

and

$$\hat{\mathbf{H}}_v(n) = \sqrt{N}\left(\mathbf{F} \otimes \mathbf{I}_{M_R M_T}\right)\left[\hat{\mathbf{h}}^T(n), 0, \cdots, 0\right]^T.$$

Since $\mathbf{F} \otimes \mathbf{I}_{M_R M_T}$ is unitary, the MSE can be expressed as

$$\begin{aligned} \text{MSE} &= E\left\{\left(\hat{\mathbf{H}}_v(n) - \mathbf{H}_v(n)\right)^H\left(\hat{\mathbf{H}}_v(n) - \mathbf{H}_v(n)\right)\right\} \\ &= NE\left\{\left(\hat{\mathbf{h}(n)} - \mathbf{h}(n)\right)^H\left(\hat{\mathbf{h}}(n) - \mathbf{h}(n)\right)\right\}. \end{aligned} \tag{4.12}$$

This implies that the estimation of frequency domain channel responses $\mathbf{H}_v(n)$ can be equivalently performed by estimating the time domain channel responses $\mathbf{h}(n)$. According to (4.4), the received signal is linear for the time domain channel response $\mathbf{h}(n)$, hence a linear MMSE estimation can be performed [88]

$$\hat{\mathbf{h}}(n) = \mathbf{Q}(n)\mathbf{r}(n), \tag{4.13}$$

where

$$\begin{aligned} \mathbf{Q}(n) &= \sqrt{\frac{\rho}{M_T}}\mathbf{R}_h\mathbf{A}^H(n)\left(\frac{\rho}{M_T}\mathbf{A}(n)\mathbf{R}_h\mathbf{A}^H(n) + \mathbf{I}_{NM_R}\right)^{-1} \\ &= \sqrt{\frac{\rho}{M_T}}\left(\mathbf{R}_h^{-1} + \frac{\rho}{M_T}\mathbf{A}^H(n)\mathbf{A}(n)\right)^{-1}\mathbf{A}^H(n). \end{aligned} \tag{4.14}$$

The error covariance is

$$\begin{aligned}\mathbf{C}_e &= E\left\{\left(\mathbf{h}(n)-\hat{\mathbf{h}}(n)\right)\left(\mathbf{h}(n)-\hat{\mathbf{h}}(n)\right)^H\right\} \\ &= \left(\mathbf{R}_h^{-1}+\frac{\rho}{M_T}\mathbf{A}^H(n)\mathbf{A}(n)\right)^{-1}. \end{aligned} \tag{4.15}$$

Note we assume that $\mathbf{R}_h$ is invertible in the above analysis. If $\mathbf{R}_h$ is not invertible, we can add $\epsilon\mathbf{I}_{M_R M_T L}$ to $\mathbf{R}_h$, then let $\epsilon \to 0$. The estimation error can be obtained by

$$\text{MSE} = N\text{tr}\left(\mathbf{C}_e\right). \tag{4.16}$$

4.2.4 Optimum Training Sequences

From the previous section, the estimation error depends on the channel correlation matrix $\mathbf{R}_h$ and training sequences through matrix $\mathbf{A}(n)$. In this section, we investigate the optimum training sequences design that minimizes the estimation error.

Denote the eigen-decomposition of transmit and receive corrlation matrix as

$$\mathbf{R}_{rl} = \mathbf{U}_l\mathbf{\Lambda}_{rl}\mathbf{U}_l^H,$$

$$\mathbf{R}_{tl} = \mathbf{V}_l\mathbf{\Lambda}_{tl}\mathbf{V}_l^H,$$

where $\mathbf{U}_l,\ \mathbf{V}_l$ are $M_R \times M_R$ and $M_T \times M_T$ unitary matrices, respectively, and $\Lambda_{rl},\ \Lambda_{tl}$ are diagnal matrices. By the property of Kronecker product,

$$\mathbf{R}_l = \left(\mathbf{U}_l \otimes \mathbf{V}_l\right)\left(\mathbf{\Lambda}_{rl} \otimes \mathbf{\Lambda}_{tl}\right)\left(\mathbf{U}_l \otimes \mathbf{V}_l\right)^H. \tag{4.17}$$

From (4.10) and (4.17), the eigen-decomposition of $\mathbf{R}_h$ can be expressed as,

$$\mathbf{R}_h = \mathbf{U}_h\mathbf{\Lambda}_h\mathbf{U}_h^H, \tag{4.18}$$

where

$$\mathbf{U}_h = \text{diag}([\mathbf{U}_0 \otimes \mathbf{V}_0, \mathbf{U}_1 \otimes \mathbf{V}_1, \cdots, \mathbf{U}_{L-1} \otimes \mathbf{V}_{L-1}]),$$

and

$$\mathbf{\Lambda}_h = \text{diag}([\mathbf{\Lambda}_{r0} \otimes \mathbf{\Lambda}_{t0}, \mathbf{\Lambda}_{r1} \otimes \mathbf{\Lambda}_{t1}, \cdots, \mathbf{\Lambda}_{r(L-1)} \otimes \mathbf{\Lambda}_{t(L-1)}]).$$

From the decomposition (4.15) and (4.18), we have

$$\begin{aligned} \text{MSE} &= N\text{tr}\left(\mathbf{C}_e\right) \\ &= N\text{tr}\left(\mathbf{U}_h^H \mathbf{C}_e \mathbf{U}_h\right) \\ &= N\text{tr}\left(\left(\mathbf{\Lambda}_h^{-1} + \frac{\rho}{M_T}\mathbf{U}_h^H \mathbf{A}^H(n)\mathbf{A}(n)\mathbf{U}_h\right)^{-1}\right), \end{aligned} \tag{4.19}$$

By the theory of majorization [86], the matrix $\mathbf{U}_h^H \mathbf{A}^H(n)\mathbf{U}_h\mathbf{A}(n)$ must be diagonal to minimize the MSE. i.e.,

$$\mathbf{U}_h^H \mathbf{A}^H(n)\mathbf{A}(n)\mathbf{U}_h = \mathbf{\Lambda}_s \tag{4.20}$$

for some diagonal matrix $\mathbf{\Lambda}_s$. This condition implies that the columns of matrix $\mathbf{A}(n)\mathbf{U}_h$ have to be orthogonal. Note that $\mathbf{U}_h$ is a block diagonal matrix, then we have

$$\mathbf{A}(n)\mathbf{U}_h = \begin{pmatrix} \mathbf{B}_0(n,0) & \mathbf{B}_1(n,0) & \cdots & \mathbf{B}_{L-1}(n,0) \\ \mathbf{B}_0(n,1) & \mathbf{B}_1(n,1) & \cdots & \mathbf{B}_{L-1}(n,1) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{B}_0(n,N-1) & \mathbf{B}_1(n,N-1) & \cdots & \mathbf{B}_{L-1}(n,N-1) \end{pmatrix},$$

where submatrix

$$\begin{aligned} \mathbf{B}_l(n,k) &= \mathbf{D}_l(n,k)\left(\mathbf{U}_l \otimes \mathbf{V}_l\right) \\ &= e^{-\jmath 2\pi \frac{kl}{N}}\left(\mathbf{s}^T(n,k)\mathbf{U}_l\right) \otimes \mathbf{V}_l. \end{aligned} \tag{4.21}$$

Let

$$\mathbf{f}_l = \frac{1}{\sqrt{N}}\left[1, e^{-\jmath 2\pi l/N}, \cdots, e^{-\jmath 2\pi (N-1)l/N}\right]^T$$

be the $(l+1)$-th column of the unitary DFT matrix $\mathbf{F}$,

$$\mathbf{S}(n) = \left[\mathbf{s}(n,1), \mathbf{s}(n,2), \cdots, \mathbf{s}(n,N)\right]^T \tag{4.22}$$

be the training signal matrix at n-th block,

$$\begin{aligned} \tilde{\mathbf{S}}(n) &= \left[\tilde{\mathbf{S}}_1(n), \tilde{\mathbf{S}}_2(n), \cdots, \tilde{\mathbf{S}}_L(n)\right] \\ &= \mathbf{S}(n)\left[\mathbf{U}_1, \mathbf{U}_2, \cdots, \mathbf{U}_L\right], \end{aligned} \tag{4.23}$$

and $\mathbf{B}_l(n) = \left[\mathbf{B}_l^T(n,1), \mathbf{B}_l(n,2), \cdots, \mathbf{B}_l^T(n,N)\right]^T$. Then $\mathbf{B}_l(n)$ can be expressed as

$$\mathbf{B}_l(n) = \sqrt{N}\left(\mathrm{diag}(\mathbf{f}_l)\tilde{\mathbf{S}}_l(n)\right) \otimes \mathbf{V}_l. \tag{4.24}$$

To satisfy equation (4.20), we have

$$\begin{aligned} \mathbf{B}_l^T(n)\mathbf{B}_l^T(n) &= N\left(\tilde{\mathbf{S}}_l^H(n)\tilde{\mathbf{S}}_l(n)\right) \otimes \left(\mathbf{V}_l^H\mathbf{V}_l\right) \\ &= N\left(\tilde{\mathbf{S}}_l^H(n)\tilde{\mathbf{S}}_l(n)\right) \otimes \mathbf{I}_{M_R} \end{aligned} \tag{4.25}$$

to be diagonal matrix, or

$$\tilde{\mathbf{S}}_l^H(n)\tilde{\mathbf{S}}_l(n) = \mathbf{U}_l^H\mathbf{S}^H(n)\mathbf{S}(n)\mathbf{U}_l, \tag{4.26}$$

be diagonal matrix. And

$$\begin{aligned} \mathbf{B}_{l'}^T(n)\mathbf{B}_l^T(n) &= N\left(\tilde{\mathbf{S}}_{l'}^H(n)\mathrm{diag}(\mathbf{f}_{l-l'})\tilde{\mathbf{S}}_l(n)\right) \otimes \left(\mathbf{V}_{l'}^H\mathbf{V}_l\right) \\ &= \mathbf{0}_{M_RM_T}, \end{aligned} \tag{4.27}$$

where $l \neq l'$. This is equivalent to

$$\begin{aligned} \tilde{\mathbf{S}}_{l'}^H(n)\mathrm{diag}(\mathbf{f}_{l-l'})\tilde{\mathbf{S}}_l(n) &= \mathbf{U}_{l'}^H\mathbf{S}^H(n)\mathrm{diag}(\mathbf{f}_{l-l'})\mathbf{S}(n)\mathbf{U}_l \\ &= \mathbf{0}_{M_T} \end{aligned} \tag{4.28}$$

or

$$\mathbf{S}^H(n)\mathrm{diag}(\mathbf{f}_{l-l'})\mathbf{S}(n) = \mathbf{0}_{M_T}. \tag{4.29}$$

Then, the conditions for the optimum training sequences become equation (4.26) and (4.29). The resulting optimum training sequences depend on the unitary matrices $\mathbf{U}_l$'s, or transmit correlation matrices $\mathbf{R}_{tl}$'s, which is determined by the angle spread of l-th scatterer cluster. If the angular spread is large, $\mathbf{R}_{tl}$ has high rank; otherwise, the rank of $\mathbf{R}_{tl}$ will decrease. In the extreme case where the angular spread is zero, $\mathbf{R}_{tl}$ degrades to rank one matrix and can be written as $\mathbf{R}_{tl} = \beta_l\mathbf{a}(\theta_{lt})\mathbf{a}^H(\theta_{lt})$ with

$\mathbf{a}(\theta_{lt})$ being the array response vector. For uniform linear array, the array response is given by

$$\mathbf{a}(\theta_{lt}) = [1, e^{\jmath 2\pi\Delta\cos(\theta_{lt})}, \cdots, e^{\jmath 2\pi\Delta(M_T-1)\cos(\theta_{lt})}]^T. \tag{4.30}$$

In the following, we will discuss the optimum training sequence design for two kinds of cases.

4.2.4.1 General Cases

In this case, we assume arbitrary transmit correlation matrices. Then Equation (4.26) is satisfied if and only if

$$\mathbf{S}^H(n)\mathbf{S}(n) = N\mathbf{I}_{M_T}. \tag{4.31}$$

To satisfy the second condition (4.29), the columns of $\mathbf{S}(n)$ have to satisfy

$$(\mathbf{s}_i(n) \odot \mathbf{s}^*_{i'}(n))^T\mathbf{f}_m = \mathbf{0}_{N\times 1}, \tag{4.32}$$

where $-L+1 \le m \le L-1$ and $\mathbf{s}_i(n), \mathbf{s}_{i'}(n)$ are i and i'-th column of $\mathbf{S}(n)$. One optimum training sequence design satisfying (4.31) and (4.32) is given at [82],

$$\mathbf{s}_0(n) \odot \mathbf{s}^*_i(n) = \sqrt{N}\mathbf{f}_{i\bar{L}},$$

where $i = 1, \cdots, M_T-1$, and $\bar{L} \ge L$. $\mathbf{s}_0(n)$ can be any vector with constant-modulus, $|s_0(n,k)| = 1$. The above optimum training sequences require that $\bar{L}(M_T-1) \le (N-L)$. Therefore the number of optimum training sequences is limited. Since $\bar{L} \ge L$, we have $M_T \le N/L$. For an OFDM system with 64 subchannels and the tap number $L = 7$, the transmit antennas can be at most 9.

Another simple training design satisfying the conditions is circularly equal-spaced and equal power sequences, which is similar as in [83]. The training sequences for different transmit antennas are transmitted through distinct subset of subchannels. Let $\Omega = \{0, 1, \cdots, N-1\}$ be set of all subchannels and $\mathcal{P}_i = \{i_0, i_1, \cdots, i_{P-1}\} \subseteq \Omega$ be

subchannel subset for the training sequence corresponding to i-th transmit antenna. To satisfy (4.32), we set $\mathcal{P}_i \cap \mathcal{P}_{i'} = 0$ for $i \neq i'$. Then, $\mathbf{s}_i(n) \odot \mathbf{s}_{i'}^*(n) = 0$. For constant-modulus and $i' = i$, (4.32) becomes

$$(\mathbf{s}_i(n) \odot \mathbf{s}_{i'}^*(n))^T \mathbf{f}_m = \frac{\sqrt{N}}{P} \sum_{p=0}^{P-1} e^{-\jmath 2\pi \frac{mip}{N}} = 0.$$

Then, i_p's have to be circularly equal-spaced, i.e., $i_p - i_0 = p\delta$ and $N = P\delta$. For this training sequence, $\delta = N/P$ must be an integer and $\delta \geq 2$. To reduce the leakage, the number of subchannels of each training sequence, P, must be as large as possible [38]. For both optimum training sequence designs, we have $\mathbf{A}^H(n)\mathbf{A}(n) = N\mathbf{I}_{LM_RM_T}$, then

$$\mathbf{Q}(n) = \sqrt{\frac{\rho}{M_T}} \mathbf{U}_h \left(\Lambda_h^{-1} + \frac{N\rho}{M_T} \mathbf{I}_{LM_TM_R} \right)^{-1} \mathbf{U}_h^H \mathbf{A}^H(n). \tag{4.33}$$

The corresponding MSE is

$$\text{MSE} = \sum_{i=0}^{LM_TM_R-1} \frac{N\lambda_i(\mathbf{R}_h)}{1 + \frac{N\rho}{M_T}\lambda_i(\mathbf{R}_h)}. \tag{4.34}$$

4.2.4.2 Special Cases

Here we investigate training design under two special environments. In the first case, there is a dominant tap (assume the first tap for convenience), then we only need to design the training according to the correlation of the dominate tap and ignore the correlations of the other taps. For the second case, all taps have the same transmit correlation matrices, i.e., $\mathbf{U}_0 = \mathbf{U}_1 = \cdots = \mathbf{U}_{L-1} = \tilde{\mathbf{U}}$. This happens when all the taps emanate from the same angle with the same angle spread. In both cases, the training design is optimized according to only one correlation matrix. The first condition (4.26) can be satisfied if the columns of $\mathbf{S}(n)\tilde{\mathbf{U}}$ are orthogonal. Let $\mathbf{S}(n)\mathbf{U} = [\sqrt{\beta_0}\tilde{\mathbf{s}}_0(n), \sqrt{\beta_1}\tilde{\mathbf{s}}_1(n), \cdots, \sqrt{\beta_{M_T}}\tilde{\mathbf{s}}_{M_T}(n)]$, where $\mathbf{s}_i(n)$'s are normalized vectors and $\sum_{i=0}^{M_T} \beta_i = NM_T$. To satisfy the second condition (4.29), we can choose $\tilde{\mathbf{s}}_i(n)$'s the same as ones for the general cases. By (4.19), the MSE can be expressed

as

$$\text{MSE} = \sum_{i=0}^{M_T-1} \sum_{j=0}^{M_R} \sum_{l=0}^{L} \frac{1}{\frac{\rho}{M_T}\beta_i + (\lambda_i(\mathbf{R}_{tl})\lambda_j(\mathbf{R}_{rl}))^{-1}}. \tag{4.35}$$

For the second case, $\mathbf{R}_{tl} = \sigma_l \mathbf{I}_{M_T},\ l > 0$. Then, we can further optimize the parameter β_i to minimize the MSE. The optimization problem can be solved through Langrange multipliers. The optimum training sequence design can be regarded as a power allocation problem. The transmitter allocates the power over the eigen-modes of the transmit correlation to minimize the estimation error. Unfortunately, there is no close-form solution in general. We shall therefore resort to the solution at high SNR and low SNR regions.

1) *High SNR Solution* At high SNR, $\rho \to +\infty$, the MSE can be approximated as

$$\text{MSE} \approx \sum_{i=0}^{M_T-1} \frac{c_i}{\frac{\rho}{M_T}\beta_i}, \tag{4.36}$$

where c_i is the number of nonzero eigenvalues corresponding to β_i. Minimizing (4.36) with respect to β_i's by using Langrange multipliers, we obtain

$$\beta_i = \frac{NM_T\sqrt{c_i}}{\sum_{k=0}^{M_T}\sqrt{c_k}}, \tag{4.37}$$

If all of the correlation matrices are full rank, the optimum training sequence design assigns equal powers, which results in the same design as in the general cases.

2) *Low SNR Solution* When SNR is small, using Taylor expansion, we have

$$\begin{aligned} &\frac{1}{\frac{\rho}{M_T}\beta_i + (\lambda_i(\mathbf{R}_{tl})\lambda_j(\mathbf{R}_{rl}))^{-1}} \\ &\quad \approx (\lambda_i(\mathbf{R}_{tl})\lambda_j(\mathbf{R}_{rl})) - (\lambda_i(\mathbf{R}_{tl})\lambda_j(\mathbf{R}_{rl}))^2 \frac{\rho}{M_T}\beta_i. \end{aligned} \tag{4.38}$$

By (4.35), minimize MSE is equivalent to maximize the following quantity

$$\Psi = \frac{\rho}{M_T} \sum_{i=0}^{M_T-1} \left(\sum_{j=0}^{M_R} \sum_{l=0}^{L} (\lambda_i(\mathbf{R}_{tl})\lambda_j(\mathbf{R}_{rl}))^2 \right) \beta_i. \tag{4.39}$$

Let

$$k = \arg\max_i \left(\sum_{j=0}^{M_R} \sum_{l=0}^{L} (\lambda_i(\mathbf{R}_{tl})\lambda_j(\mathbf{R}_{rl}))^2 \right), \tag{4.40}$$

then the solution for β_i is,

$$\beta_i = \begin{cases} NM_T, & i = k \\ 0, & i \neq k \end{cases}. \tag{4.41}$$

That implies that the optimum training sequence design assigns the total power to one best favorable eigen-space when SNR is low.

In the above discussion, we have developed the optimum design of training sequences. With the knowledge of the spatial correlation of MIMO broadband channel, we can design training sequences to minimize channel estimation error of broadband MIMO channels. Note the estimation we obtained is based on only one OFDM block. If multiple training blocks or decision-directed estimation are used, more accurate estimation can be obtained by exploiting the time correlation of channel parameters using the method developed in [38].

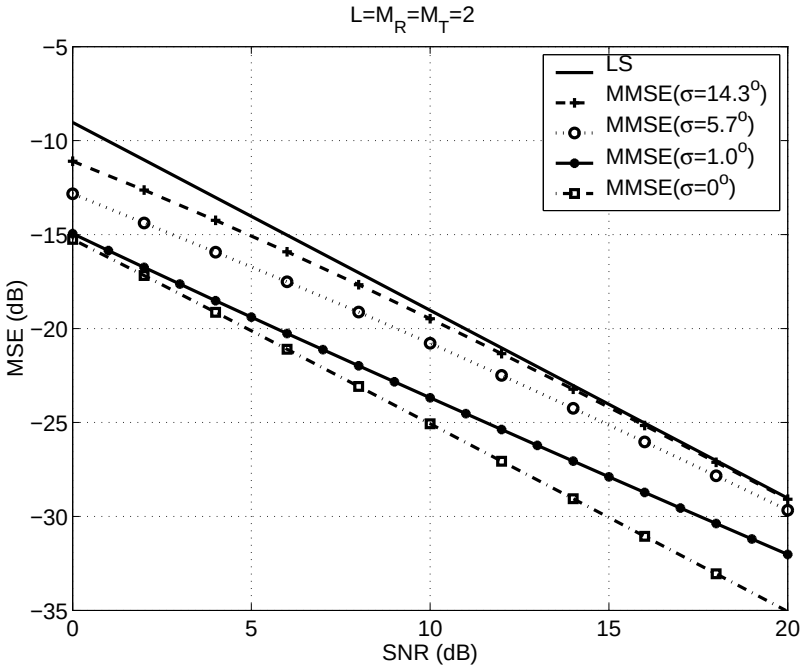


Figure 4.2: MSE of OFDM system with $L = M_T = M_R = 2$.

4.2.5 Simulation Results

In this section, we present our simulation results to demonstrate the performance of the proposed channel estimation approaches and the training sequence designs. To study the impact of the propagation parameters, we use a simple channel model in Examples 1 and 2. In the simulated system, two transmit antennas and two receive antennas are used, that is, $M_R = M_T = 2$. The entire channel is divided into 32 subchannels, and the cyclic prefix is longer than the delay spread. The channel has two taps with the power $\sigma_0^2 = 0.8$ and $\sigma_1^2 = 0.2$, respectively. The relative antenna space is $\Delta = 1$.

4.2.5.1 Simulation Example 1

In first example, we study the impact of the angular spread on the performance of the estimation. The mean angles of departure are $\bar{\theta}_{t0} = 13.5°$, $\bar{\theta}_{t1} = 26.4°$ and the angles of arrival are $\bar{\theta}_{r0} = 290.3°$, $\bar{\theta}_{r1} = 332.3°$. The angle spread for all tap at both transmitter and receiver are same.

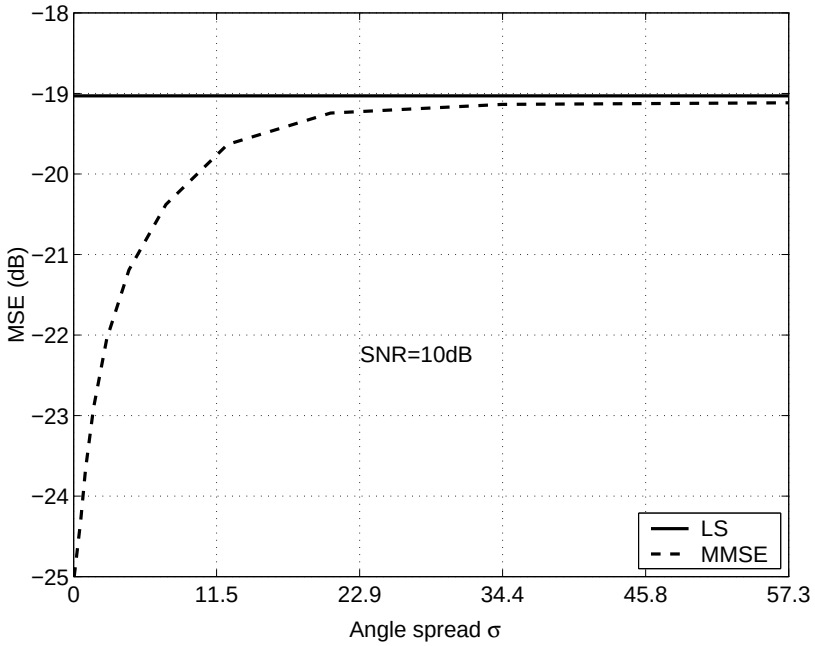


Figure 4.3: MSE vs. angle spread at a 10 dB SNR.

Figure 4.2 demonstrates the estimation MSE for different angle spread. For comparison, we also give the MSE of LS estimation. From the figure, the performance of the MMSE estimator is always better than the LS estimator. However, the improvement diminishes with the increase of SNR for nonzero angle spread. This effect can be derived from (4.15). When SNR approach infinity, the error covariance is close to the error covariance of LS estimation. We also observe that the MSE is decreasing when the angle spread decreases. That shows that the MMSE estimator can capture the spatial correlations of MIMO channels. While the performance of LS estimator is independent of the angle spread. The impact of angle spread can be seen clearly in Figure 4.3. When the angle spread is equal to zero, the correlation matrices have rank one, the total parameters to be estimated decrease from 8 to 2;, therefore a 6 dB improvement is obtained for all SNR.

4.2.5.2 Simulation Example 2

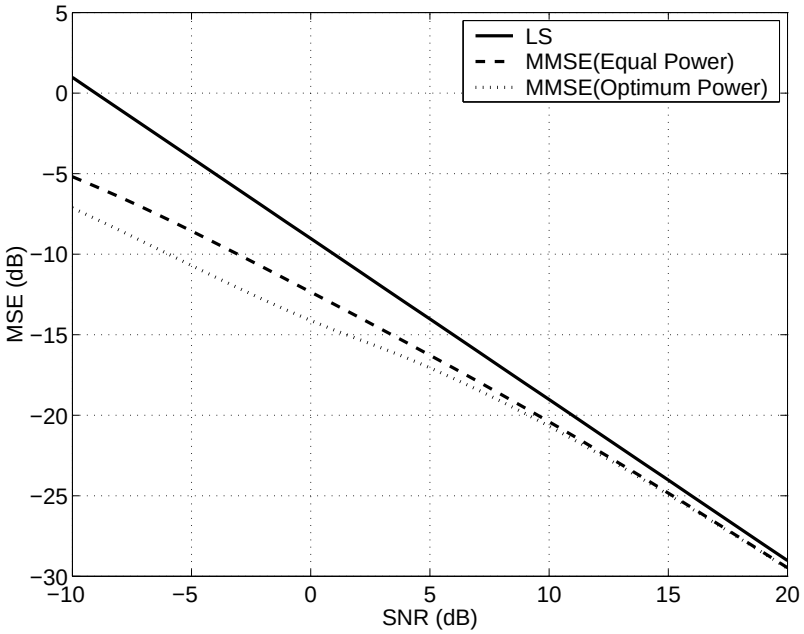


Figure 4.4: MSE of estimation for the MIMO channel with the same angle of departure.

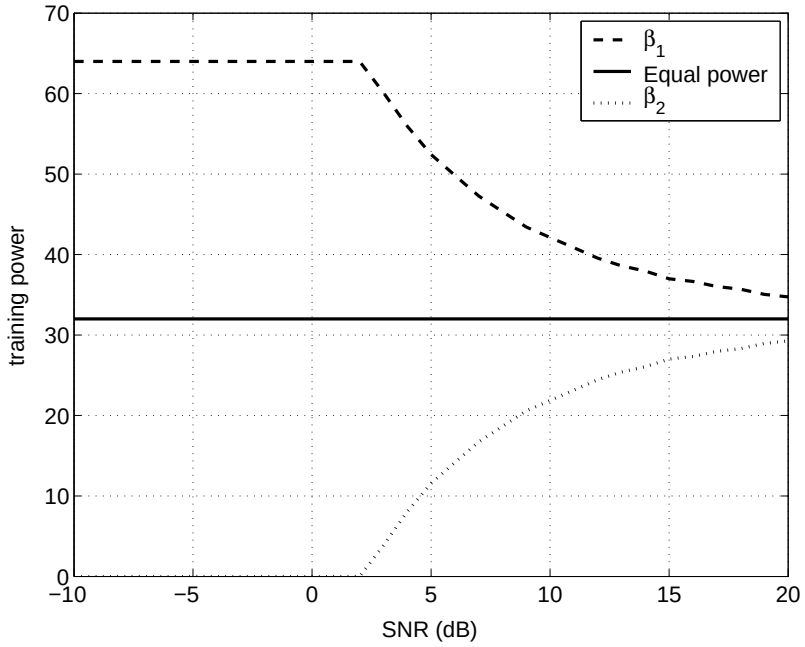


Figure 4.5: Optimum power allocation of the training sequences.

In this example, we set the angle of departure for two taps to be same. The angle spread is 0.15 rad. The other parameters are the same as in Example 1. Then, we can use the training design derived for the special cases. Figure 4.4 shows the performance of estimation using LS, MMSE with equal power training and MMSE with optimum power allocation. From the figure, the training with optimum power allocation achieves the best performance, especially at low SNR region. Figure 4.5 gives the optimal power distribution. At low SNR, the optimum training tends to assign all the power to one favorable branch. When SNR increase, the optimum training assigns the power equally over all eigen-spaces. Therefore, at the high SNR region, the performance of three approaches are close.

4.2.5.3 Simulation Example 3

In this example, we use the indoor MIMO WLAN channel model case C proposed by the IEEE 802.11 TGn channel model special committee [93]. The MIMO-OFDM system parameters follow the parameters of IEEE Std 802.11a. The entire bandwidth,

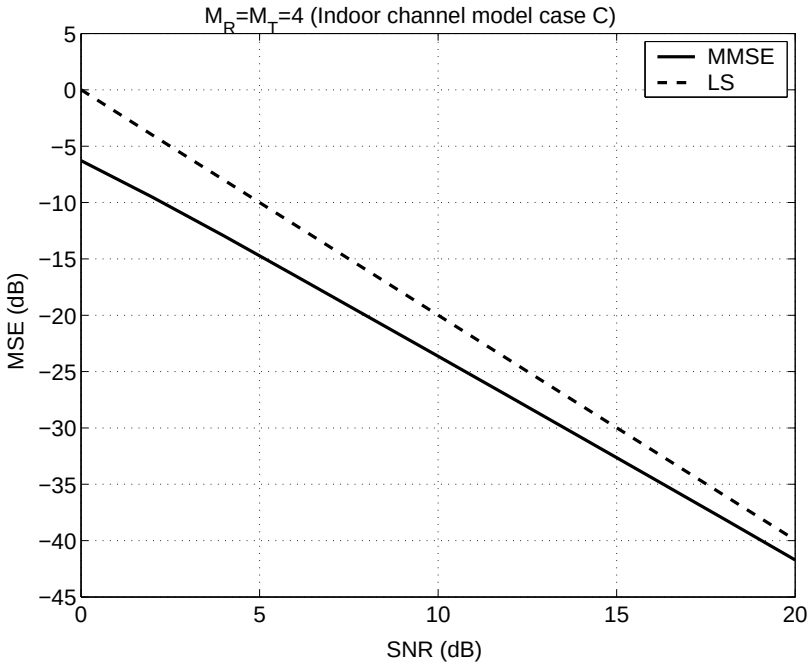


Figure 4.6: MSE of estimation for a 4×4 system.

20 MHz, is divided into 64 sub-channels with subchannel space $\Delta f = 0.3125$MHz. The symbol duration is $T_s = 3.2$ μs. An additional 0.8 μs guard interval is used to mitigate inter-symbol interference. Four transmit and four receive antennas are used. The Matlab program that generates the MIMO channels was written by L. Schumacher [94]. Since the channel correlation matrices of all taps are non-singular, we use the training sequences given in Section 4.2.4.1. The MSEs per subchannel for both MMSE and LS channel estimations are presented in Figure 4.6. From the simulation results, there are $2 \sim 6$ dB performance gain for our approach than the LS approach. As mentioned previously, the performance of both approaches tends to be close at high SNR region.

4.3 Transmission with Channel Feedback

When the transmitter does not have any knowledge of the channel state information. Therefore, the signal transmission will be independent of the channel information.

However, if either full or partial channel information is available at the transmitter, both diversity and multiplexing gain can be obtained by exploiting the channel state information with low complexity. It has been proved that the capacity and outage probability of a MIMO system can be substantially improved if partial channel state information is available at the transmitter [7], [9].

The channel knowledge can be obtained either through feedback from the receiver or exploiting reciprocity principle in a duplex system. Depending on the applications, different levels of accuracy in CSI are needed. In single antenna systems, power and constellation size of the modulation can be adaptively adjusted according to the channel amplitude so that the capacity or outage probability is maximized. For MIMO systems, both amplitude and phase of the channel are important and the problem is more complicated. The schemes of exploiting channel information are determined by a variety of factors such as the nature of channel information, power constraints, choice of signaling and coding, receiver and performance criteria.

If full channel information is available at the transmitter, the optimal scheme that maximizes the capacity is to apply a linear precoding matrix to the transmitted signal. Let $\mathbf{H}$ be the MIMO channel matrix with singular value decomposition $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$. The optimal precoding matrix can expressed as $\mathbf{V}\hat{\mathbf{\Sigma}}^{1/2}$, where $\hat{\mathbf{\Sigma}}$ is and diagonal matrix with the diagonal elements obtained by water-pouring based power allocation [5].

In practice, assumption of the full channel information at the transmitter appears to be unrealistic in most of situations because only limited feedback overhead is allowed. In some communication systems such as cellular downlink channel with multiple antennas installed at the base station, the channel experiences transmission correlation and the correlation matrix is known to the transmitter by either implicit or explicit feedback. Since the transmission correlation is relatively stable compared with the instantaneous channel matrix, only small feedback overhead is needed. For

the MIMO channel with transmission correlation, the channel matrix can be modelled as $\mathbf{H} = \mathbf{H}_w \mathbf{R}_t^{1/2}$, where $\mathbf{H}_w$ is a random matrix and $\mathbf{R}_t$ is the correlation matrix. The optimal linear precoding matrix is similar to that of full channel information case [10].

If the transmit provides more antennas than the transmitted substreams, the error performance can be improved by transmit diversity. Linear precoding, which spreads the signal over all transmit antennas using a precoding matrix, is a simple and efficient way to achieve transmit diversity [7]. However, optimal precoding requires accurate knowledge of channel information. Sometimes, we can only feedback partial information. In this case, we propose a precoding approach. The basic idea is to quantize the MIMO channel using a pre-designed linear precoding codebook. Both transmitter and receiver know the codebook. The receiver selects the best codeword according to some performance criteria and convey the index of the codewords. The codebook can be designed to achieve the desired performance and feedback bandwidth tradeoff. For indoor MIMO-OFDM systems, the feedback bits can be further reduced by exploiting the channel correlations at the frequency domain. In [101], an interpolation-based beamforming approach is proposed to reduce the feedback requirement. However, this approach works only for single data stream transmission and cannot be extended to multiple data streams transmission. In this section, we describe a new approach that is robust to any MIMO systems and works well for multiple data stream transmission.

4.3.1 Precoded MIMO-OFDM System

The diagram for a precoded MIMO-OFDM system with M_T transmit antennas and M_R (assume $M_T \geq M_R$) receive antennas and N subcarriers is given by Figure 4.7. OFDM converts the broadband frequency selective channel into N narrowband flat fading subchannels provided the cyclic prefix is longer than the channel delay spread.

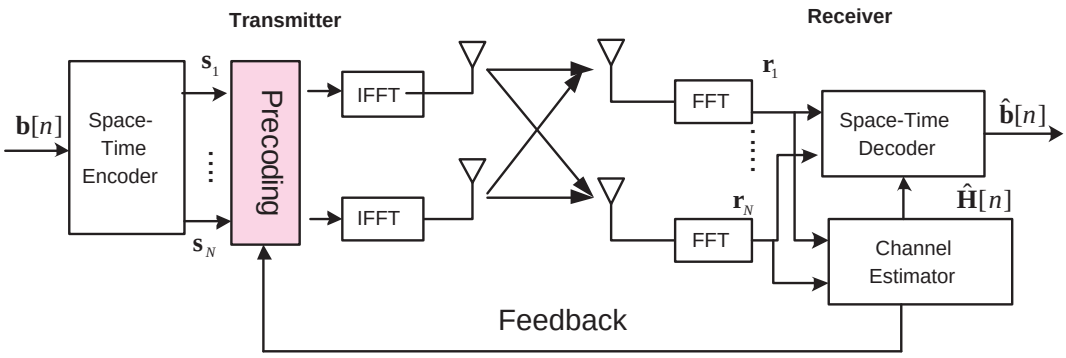


Figure 4.7: Block diagram of a MIMO-OFDM system with precoding.

For each subchannel, M_R data substreams are modulated and multiplied by the precoding matrices $\mathbf{P}_i$'s to form the signals transmitted through the antennas. The received signal at one subchannel can be written

$$\mathbf{r}_k = \mathbf{H}_k \mathbf{P}_{i_k} \mathbf{s}_k + \mathbf{n}_k, \tag{4.42}$$

where $\mathbf{H}_k \in \mathbb{C}^{M_R \times M_T}$ is the random channel matrix, $\mathbf{s}_k$ is the transmitted signal with $E\{\mathbf{s}_k \mathbf{s}_k\} = \frac{\mathcal{E}_s}{M_R} \mathbf{I}_{M_R}$, $\mathbf{P}_{i_k}$ is a $M_T \times M_R$ precoding matrix, and $\mathbf{n}_k$ is AWGN with variance N_0. To maintain the power of the transmitted signal, we restrict the precoding matrices to be unitary, i.e., $\mathbf{P}_i \in \mathcal{U}_{M_T, M_R}$ with $\mathcal{U}_{M_T, M_R}$ being the set of all $M_T \times M_R$ complex matrices with orthonormal columns. The receiver is assumed to have perfect channel state information. Since the feedback channel capacity is limited, the precoding matrices have to be quantized using a predetermined codebook $\mathcal{P} = \{\mathbf{P}_1, \cdots, \mathbf{P}_L\}$. The receiver picks up a codeword from the codebook $\mathcal{P}$ according to some performance optimization criteria and sends the indices of the codeword to the transmitter. Our interest focuses on the codebook design, optimal codeword selection and feedback bit reduction for OFDM systems. We present several techniques to solve these problems in the following sections.

4.3.2 Codebook Construction

In this section, we first describe the problem of codebook design. A codebook of size L is a set of $M_T \times M_R$ unitary matrices $\mathcal{P} = \{\mathbf{P}_1, \cdots, \mathbf{P}_L\}$. It is shown [100] that the codebook should be designed as a set of subspaces rather than as a set of matrices, i.e., the codebook is a finite set of points on the Grassmann manifold G_{M_T,M_R} [104], which is defined as the set of all M_R-dimensional subspaces of an M_T-dimensional vector space $\mathbb{C}^{M_T}$ (complex dimension). A point on Grassmann manifold is a linear subspace, which can be specified by an arbitrary orthogonal basis stored as an $M_T \times M_R$ unitary matrix.

The chordal distance in G_{M_T,M_R} is defined as

$$d_c^2(\mathbf{P}_l, \mathbf{P}_{l'}) = \sum_{m=1}^{M} \sin^2(\theta_m) \tag{4.43}$$

$$= M - \|\mathbf{P}_l^H \mathbf{P}_{l'}\|^2, \tag{4.44}$$

where θ_m's are the principle angles between the subspaces spanned by the columns of $\mathbf{P}_l$ and $\mathbf{P}_{l'}$, and $\|\cdot\|$ is the Frobenius norm.

If the entries of the channel matrix $\mathbf{H}$ are independently Gaussian distributed, the subspaces spanned by the eigenvectors corresponding to the maximal eigenvalues of the channel are isotropically distributed [100]. Then the codebook design becomes to a subspace packing in Grassmann manifold with large minimum distance. In general, the optimal packing is difficult or impossible to find [102]. Therefore, some suboptimal approaches for codebook design with large minimum distance are developed for noncoherent constellation design [103], [98]. Most of design constraints the structure of the codebook to reduce the search area.

A simple and straightforward method to find the codebook with the required minimum distance is totally random search. Assume we have already a set of matrices $\{\mathbf{P}_0, \cdots, \mathbf{P}_{i-1}\}$ with minimum distance larger than a predetermined value d_0. Each

step, the algorithm randomly generate $M_T \times M_R$ matrix $\mathbf{H}_i$ with elements independently Gaussian distributed. From the random matrix $\mathbf{H}_i$, we obtain the unitary matrix $\mathbf{P}_i$ with its columns spanning the same subspace as that spanned by the columns of $\mathbf{H}_i$. Calculate the minimum distance between $\mathbf{P}_i$ and $\{\mathbf{P}_0, \cdots, \mathbf{P}_{i-1}\}$. If the minimum distance is greater than d_0, add $\{\mathbf{P}_i$ into $\{\mathbf{P}_0, \cdots, \mathbf{P}_{i-1}\}$. Otherwise, discard it and start with a new $\mathbf{H}_i$. The procedure continues till all required codewords are found. When the codebook size is small, the random search approach is very fast and can obtain a good codebook with a large minimum distance.

4.3.3 Precoding Matrix Selection Criteria

In this section, we introduce several selection criteria that can be used to select the precoding codeword to optimize the performance.

4.3.3.1 Maximize the Minimum Distance for Maximum Likelihood Receiver

ML detector is an optimal detector that minimizes symbol error rate. An ML solves the following problem

$$\hat{\mathbf{s}}_k = \arg\min_{\mathbf{s}_k} \|\mathbf{r}_k - \mathbf{H}_k \mathbf{P}_{i_k} \mathbf{s}_k\|, \tag{4.45}$$

where $\|\cdot\|$ is Euclidean norm. Usually, the error probability of ML detector can be upper bounded by the union bound that is determined by the minimum distance. Therefore, the selection criterion for ML detector is to select the precoding matrix so that the minimum distance $d_{min}(\mathbf{H})$ of the precoded signal is maximized. For a channel matrix $\mathbf{H}$, we choose the precoding matrix as

$$\mathbf{P}^* = \arg\max_{\mathbf{P}_i \in \mathcal{P}} d^2_{min}(\mathbf{H}, \mathbf{P}_i) \tag{4.46}$$

Assume that the minimum distance is achieved between two signal vectors $\tilde{\mathbf{S}} = \mathbf{s}_i - \mathbf{s}_j$. Then the minimum distance at the receiver is

$$d^2_{min}(\mathbf{H}) = \|\mathbf{H}\mathbf{P}\tilde{\mathbf{s}}\|^2, \tag{4.47}$$

and the optimal precoding codeword is

$$\mathbf{P}^* = \mathbf{P}_l,\ l = \arg \max_{l' \in \{1,\cdots,L\}} d^2_{min}(\mathbf{H}). \tag{4.48}$$

The maximization is over all possible precoded signal difference vectors and requires high computational complexity. An sub-optimal method can be obtained by using Cauchy-Schwarz inequality

$$d^2_{min}(\mathbf{H}) = \|\mathbf{HP}\tilde{\mathbf{S}}\|^2 \leq \|\mathbf{HP}\|^2 \|\tilde{\mathbf{S}}\|^2. \tag{4.49}$$

Then, the selection criterion becomes to select the $\mathbf{P}$ that maximizes the Frobenius norm of the equivalent channel matrix, $\|\mathbf{HP}\|$. The computational complexity is then reduced significantly.

4.3.3.2 Minimize MSE for Zero-Forcing and MMSE Receiver

For zero-forcing and MMSE receivers, we choose the precoding codeword to minimize the MSE. The MSE's of two linear receivers are

$$\text{MSE}_{ZF} \quad = \quad N_0 \text{tr}\left\{(\mathbf{P}^H \mathbf{H}^H \mathbf{HP})^{-1}\right\}, \tag{4.50}$$

and

$$\text{MSE}_{MMSE} \quad = \quad \frac{E_s}{N_0} \text{tr}\left\{\left(\mathbf{I}_{M_T} + \frac{E_s}{M_R N_0} \mathbf{P}^H \mathbf{H}^H \mathbf{HP}\right)^{-1}\right\}. \tag{4.51}$$

The optimal precoding matrix is

$$\mathbf{P}^* = \arg \max_{\mathbf{P}_l \in \mathcal{P}} \text{MSE}_{ZF}(\text{MSE}_{MMSE}). \tag{4.52}$$

Let the singular value decomposition is

$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H.$$

It has been shown [100] that the optimal precoding matrix is

$$\mathbf{P}_{opt} = \mathbf{V}', \tag{4.53}$$

where $\mathbf{V}'$ is the first M column of $\mathbf{V}$. As an suboptimal way, we can select $\mathbf{P}_l$ that is closest to the optimal precoding matrix $\mathbf{P}_{opt}$,

$$\mathbf{P}^* = \arg \min_{\mathbf{P}_i \in \mathcal{P}} d_c(\mathbf{P}_i, \mathbf{V}'). \tag{4.54}$$

The suboptimal approach does not need to calculate the matrix inversion, therefore, the complexity is reduced.

4.3.3.3 Maximize the Instantaneous Capacity

The instantaneous capacity of a precoded system with channel matrix $\mathbf{H}$ is given by

$$C(\mathbf{P}_l) = \log \det \left(\mathbf{I}_{M_R} + \frac{E_s}{M_R N_0} \mathbf{P}_l^H \mathbf{H}^H \mathbf{H} \mathbf{P}_l \right). \tag{4.55}$$

The optimal precoding matrix that maximizes the capacity is shown to be $\mathbf{P}_{opt} = \mathbf{V}'$, then we can use the same suboptimal selection rule as that of previous section (4.54). At low SNR, the capacity of can be approximated as

$$C(\mathbf{P}_l) \approx \frac{E_s}{M_R N_0} \mathrm{Tr} \left\{ \mathbf{P}_l^H \mathbf{H}^H \mathbf{H} \mathbf{P}_l \right\}. \tag{4.56}$$

Then, we can choose the precoding codeword by

$$\mathbf{P}^* = \mathbf{P}_l, \ l = \arg \max_{l' \in \{1, \cdots, L\}} \mathrm{tr} \left\{ \mathbf{P}_l^H \mathbf{H}^H \mathbf{H} \mathbf{P}_l \right\}. \tag{4.57}$$

4.3.4 Precoded MIMO-OFDM System with Limited Feedback

In the above sections, we have considered precoding only for narrow band system. The precoding method can be directly extended to a broadband MIMO system using OFDM modulation. For the direct extension of the precoding method, the receiver selects the optimal precoding matrix from the same codebook for each subchannel. Therefore, the feedback bits is proportional to the number of the subcarriers of the OFDM system. If the codebook size is L and there are N subcarriers, we need to use $N \log_2 L$ feedback bits totally.

It is observed that the channels responses for the subchannels in an OFDM system are highly correlated [38]. The correlation is fully determined by the channel delay spread and independent of the number of the subcarrier. Then, we can make use of the channel correlation to reduce the feedback bit rate. However, we cannot directly use the correlation because the optimal precoding matrix contains only the subspace information of the MIMO channels. While the exact channel information has not only subspace information but the amplitude of the MIMO channel at the subspaces [54]. One method to exploit the channel correlation is to feedback both channel subspace and amplitude information. This is the approach by Choi and Heath [101]. However, this approach is only applicable for single data stream transmission as shown in [101].

We propose to use a method based on the idea of subspace tracking. In the proposed method, only channel subspace information is fed back. To reduce the feedback bits, we select the precoding codewords from a subset of the codebook, therefore at each step, the feedback bit rate can be reduced significantly. We found this approach is robust to multiple data stream transmission.

4.3.4.1 Interpolation Based Beamforming

In this section, we introduce the interpolation-based approach [101]. To reduce the feedback bit rate, we combine the K adjacent subcarriers into a cluster, and each cluster uses one precoding matrix. This method is called clustering method. Using clustering method, the feedback rate can be reduced by $1/K$. The disadvantages of clustering method is that the performance at the boundaries of the cluster is degraded significantly. As an alternative, interpolation approach can improve the performance with a small increased feedback bits. The proposed interpolation approach is described as follows. Assume the precoding matrices of two adjacent clusters are $\mathbf{P}_{lK+1}$ and $\mathbf{P}_{(l+1)K+1}$, then the precoding matrices at the subcarriers between the two clusters

can be obtained through following interpolation formula,

$$\hat{\mathbf{P}}_{lK+k;\theta_l} = \frac{(1-\lambda_k)\mathbf{P}_{lK+1} + \lambda_k e^{\jmath\theta_l}\mathbf{P}_{(l+1)K+1}}{\|(1-\lambda_k)\mathbf{P}_{lK+1} + \lambda_k e^{\jmath\theta_l}\mathbf{P}_{(l+1)K+1}\|}, \tag{4.58}$$

where $\lambda_k = (k-1)/K$ is a linear weight, θ_l is a phase rotation, which gives part of information of the channel complex amplitude. With the phase θ_l, the distortion caused by the amplitude ambiguous can be partially corrected.

To optimize the interpolation performance, the receiver quantize the phase and select the best one to send back to the transmitter as well as the indices of the precoding matrices. Since there is only one data stream, the optimization criterion is simplified to maximize the channel gain. To minimize the error rate, we only need to maximize the minimal channel gain. That is

$$\theta_l = \arg\max_{\theta\in\Theta}\{\|\mathbf{H}(lk+k)\hat{\mathbf{P}}(lK+k;\theta)\|,\ 1 \le l \le K\} \tag{4.59}$$

The phase can also be optimized with respect to other criteria, such as capacity.

The drawback of interpolation-based approach is that it is applicable only for single data stream transmission since the both subspace and amplitude of the channel should be conveyed through feedback channel. For multiple data stream transmission, the amplitude information requires much more feedback bits to achieve good performance. This can be explained as follows.

For a $M_T \times M_R$ MIMO system, the random channel matrix $\mathbf{H} \in \mathbb{C}^{M_R\times M_T}$ can be regarded as to belong to a vector space $\mathbb{C}^{M_T M_R}$ with dimension $M_T M_R$. We can change the coordinate system as $\mathbb{C}^{M_R\times M_T} \rightarrow \mathbb{C}^{M_R\times M_R} \times G_{M_T,M_R}$. The Grassmann manifold contains the information of channel subspace and has dimension of $M_R(M_T - M_R)$. Consider MIMO systems with 4 transmit antennas, $M_T = 4$. For single data stream transmission $M_R = 1$, the dimension of channel subspace information is $M_R(M_T - M_R) = 3$ and the channel amplitude information has dimension $M_R^2 = 1$. Therefore, only a few bits is needed to represent the amplitude information. However, for multiple data transmission, for example $M_R = 2$, the subspace information has

dimension $M_R(M_T - M_R) = 4$ and the amplitude information has dimension $M_R^2 = 4$. Then, to represent the amplitude information, we need to use the same amount of bits as to represent the subspace information. Consequently, the interpolation approach is hard to extend to the general case where multiple data streams are transmitted.

4.3.4.2 Tracking Based Precoding

To overcome the drawback of the interpolation approach, we propose a subspace tracking based approach that is robust to multiple data transmission. For an OFDM system, the channel responses at the adjacent subchannels are highly correlated, therefore, there is high correlation between the precoding matrices. However, the correlation of subspace is different from the exact channel correlation because the precoding matrices are points on the Grassmann manifold. We cannot use the linear interpolation directly to obtain the precoding matrices of other subchannels. There are many possible paths between precoding matrices corresponding to two correlated subchannels. For general subspace tracking method, we have to pick up a step size and quantize the orientation of the subspace evolution, which complicates the problem. Here we propose a simplified version of subspace tracking where precoding matrices are chosen to be dependent across the subchannels.

The idea originates from the fact that due to the statistical correlation between two neighboring subchannels, it is highly likely that the two desired precoding matrices reside within a small neighborhood in the high-dimensional Grassmann manifold.

Let $\mathcal{P}$ be the precoding codebook with size $L = 2^q$, and for each codewords $\mathbf{P}_i$, we assign a small neighborhood subset of $\mathcal{W}_i \subset \mathcal{P}$ with size $L' = 2^{q'}$, where $q' < q$. The subset $\mathcal{W}_i$ contains all L' matrices of the codebook that are closest to $\mathbf{P}_i$. We also define a step size K that is determined by the channel statistics. The precoding matrix is updated for every K subchannels. In the tracking algorithm, we start with the first K subchannels and use the full precision ($q = \log_2 L$ bits) to select one

of the best precoding matrix, $\mathbf{P}_0$, out of L possibilities. Observing that the best precoding matrix for the second step, $\mathbf{P}_1$, lies in a small neighborhood of $\mathbf{P}_0$, i.e., $\mathcal{W}_0$, we are able to narrow down our search. Since $q' < q$, we effectively reduce the number of feedback bits needed for the second step to q'. Recursively repeating this process to cover all N subchannels involved, and we end up with a total requirement of $q + (N/K - 1)q'$ feedback bits, which is much less than the $N * q$ bits necessary for the non-tracking approach. The search for $\mathbf{P}_1$ in the neighborhood of $\mathbf{P}_0$ is illustrated in Figure 4.8.

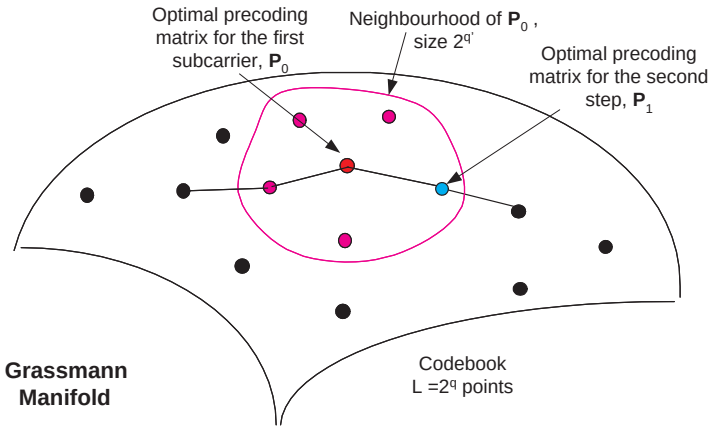


Figure 4.8: Illustration of subspace tracking in Grassmann Manifold.

We summarize the subspace tracking precoding scheme as follows. Note here we assume that the codebook is the for all the subchannels.

1. *Set step size K and use the full q bits to select the optimal precoding matrix $\mathbf{P}_0$ for the first K subchannels.*

2. *Select the best precoding matrix $\mathbf{P}_{i+1}$ from the neighborhood subset of previous precoding matrix ($\mathcal{P}_i$) for next K subchannels.*

3. *Repeat step 2 for all subchannels.*

This is a very simple and efficient approach for low bit feedback precoding. It is easy to achieve the best performance/feedback bandwidth tradeoff by selecting the step

size K and neighborhood subset size q'. The size of subset $\mathcal{W}_i$ can be determined according to the dimension of the Grassmann manifolds G_{M_T,M_R}. Heuristically, we can choose $L' = 1 \sim 2(2M_R(M_T - M_R))$.

4.3.5 Simulation Results

In this section, we present simulation results. In the simulation, the entire channel is divided into 64 subcarriers. We employ discrete channel model that has 6 taps with uniform profile and i.i.d. complex Gaussian distribution with zero-mean. The cyclic prefix is longer than the delay profile. Assume that the feedback channel has no delay and no error. The receiver knows the perfect channel information and uses MMSE detection.

We first look at the performance for narrow band MIMO systems. Figure 4.9 gives the BER for a MIMO system with 4 transmit antennas and 2 receive antennas. Two data streams are transmitted using BPSK modulation. The BER of 2×2 system without precoding is also given for comparison. From the figure, we can see that the precoded system has much better BER performance than uncoded system. For MMSE criterion, the precoded system provides about a 4 dB gain. And the slope of the BER curves are the same as the uncoded system with ML receiver. For ML precoded system, there is about a 4 dB gain over the MMSE precoded system. We also give the performance for different feedback bits. Using 4 8 bits quantization of precoding matrices, the performance is very close to the optimal one.

Figures 4.10∼4.14 address the BER performance of MIMO-OFDM system with different antennas for uncoded QPSK modulation. The precoding matrices are quantized by 6 bits. In ideal case, the receiver sends back the indices of the precoding matrix at each subcarrier and there are 384 feedback bits for one OFDM symbol. Using clustering approach, the receiver sends back one precoding matrix for each K adjacent subcarriers. In our simulation, we have performed the clustering approach

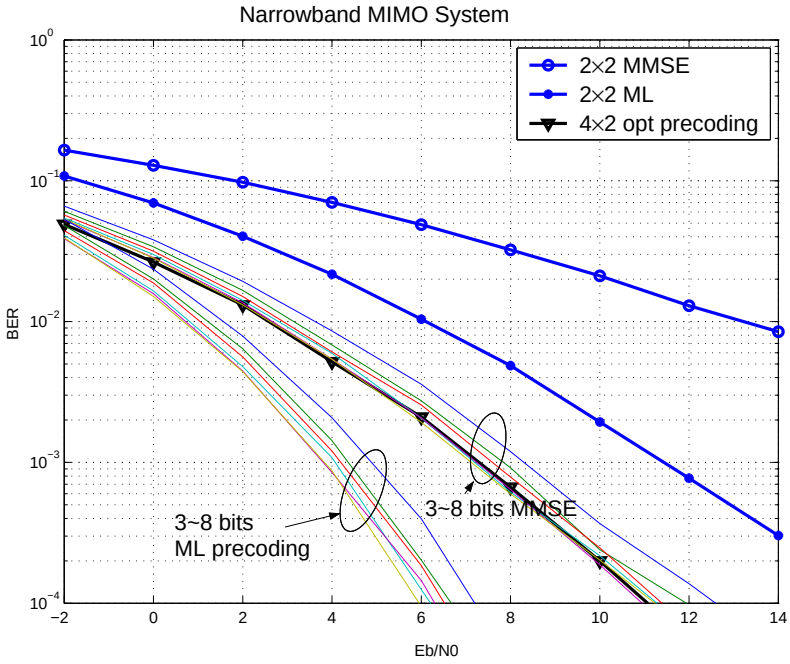


Figure 4.9: Narrow band MIMO system with precoding.

for $K = 4$ and 8, which means 96 bits and 48 bits feedback. In the tracking based approach, the precoding matrix is updated every 4 subcarriers and there are 4 bits for each step. Then plus 6 bits for the initial precoding matrix, total 66 bits are assigned. For comparison, we also present the performance of selection diversity. The selection diversity means to select the best transmit antennas within all transmit antennas. In Figures 4.10 and 4.11, the interpolation approach uses 64 feedback bits [101]. The subsampling rate is K=8, and 2 bits quantization of the rotation phase is used. From Figures 4.10 and 4.11, the interpolation (64 bits), tracking (66 bits) and clustering (96 bits) have the similar performance. The selection diversity approach is worse than the three approaches at low SNR. While the interpolation and tracking approaches have much lower feedback rates , however, their performance are very similar. From the figure, 48 bits clustering has the worst performance.

Figures 4.12∼4.14 show the performance of a precoded MIMO-OFDM system with multiple data stream transmission. In these systems, the interpolation approach

fails. From these figures, the tracking approach still works well in different systems with low feedback bits, where the selection and low rate clustering approaches suffer from performance degradation.

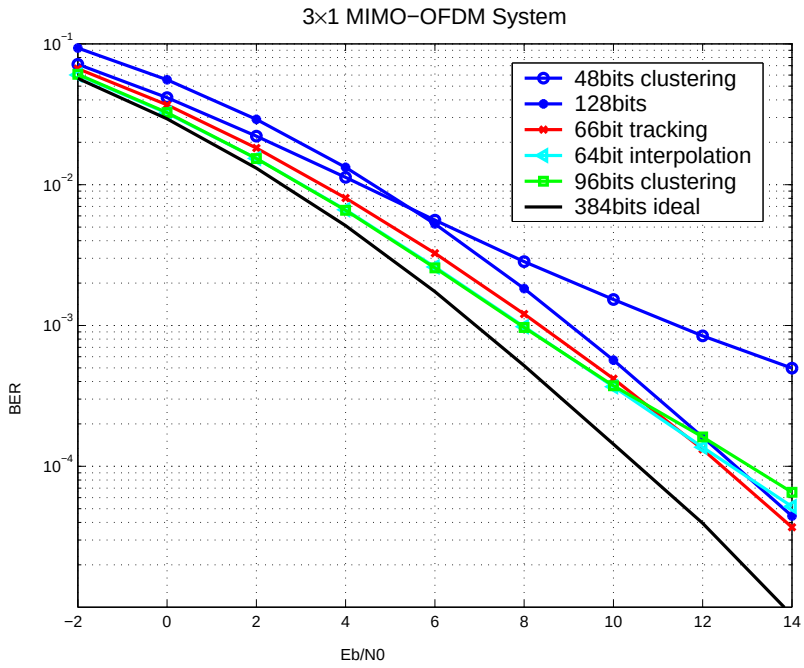


Figure 4.10: BER of a 3×1 MIMO-OFDM system with precoding.

4.4 Summary

In this chapter, we address channel estimation for MIMO-OFDM systems in spatially correlated fading channels. Exploiting the spatial correlation, the proposed channel estimator can achieve better performance than the LS estimator. Based on the MMSE estimation, we derive the conditions for the optimum training sequences and provide two optimum training sequence design approaches for arbitrary transmit correlation. The proposed approaches can be used in wireless LAN where MIMO channels are correlated.

Exploiting partial channel information can significantly improve the error performance of MIMO systems. For MIMO-OFDM systems, the previous work is based on

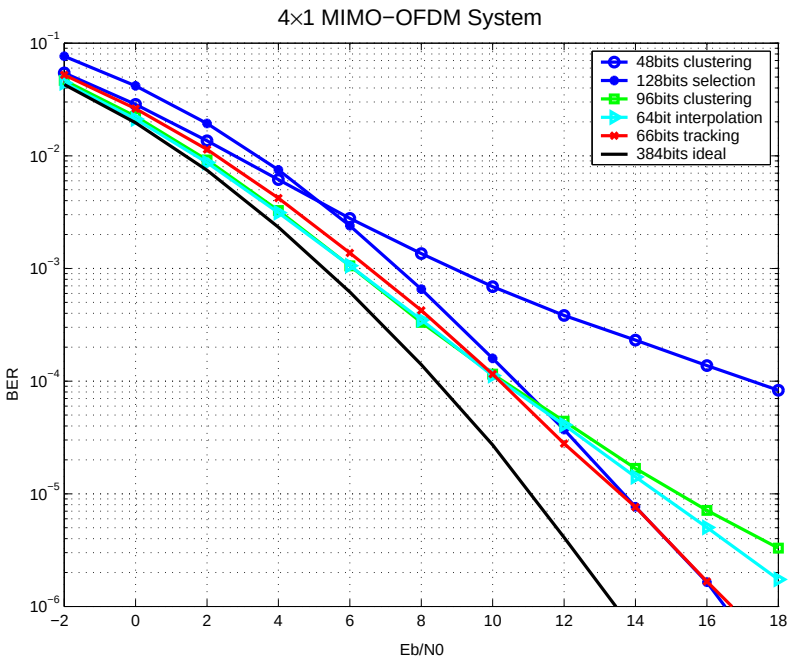


Figure 4.11: BER of a 4×1 MIMO-OFDM system with precoding.

interpolation that is designed only for single data stream transmission. Therefore, we propose a subspace tracking based approach that is robust to multiple data stream transmission. The simulation results show that the subspace tracking based approach can reduce the feedback bit rate significantly.

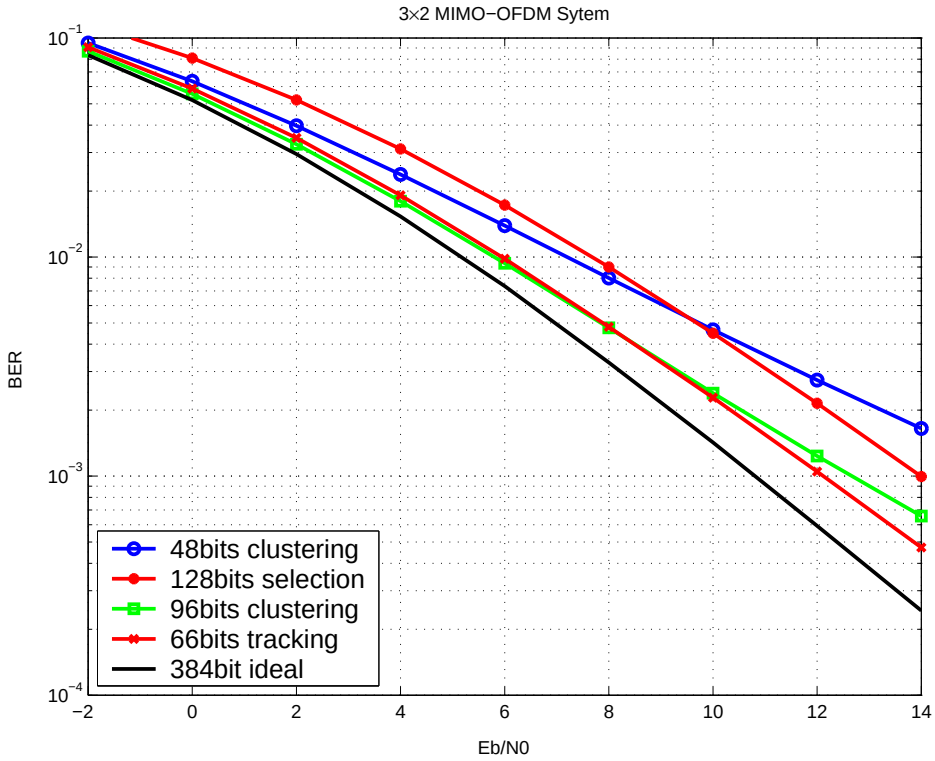


Figure 4.12: BER of a 3×2 MIMO-OFDM system with precoding.

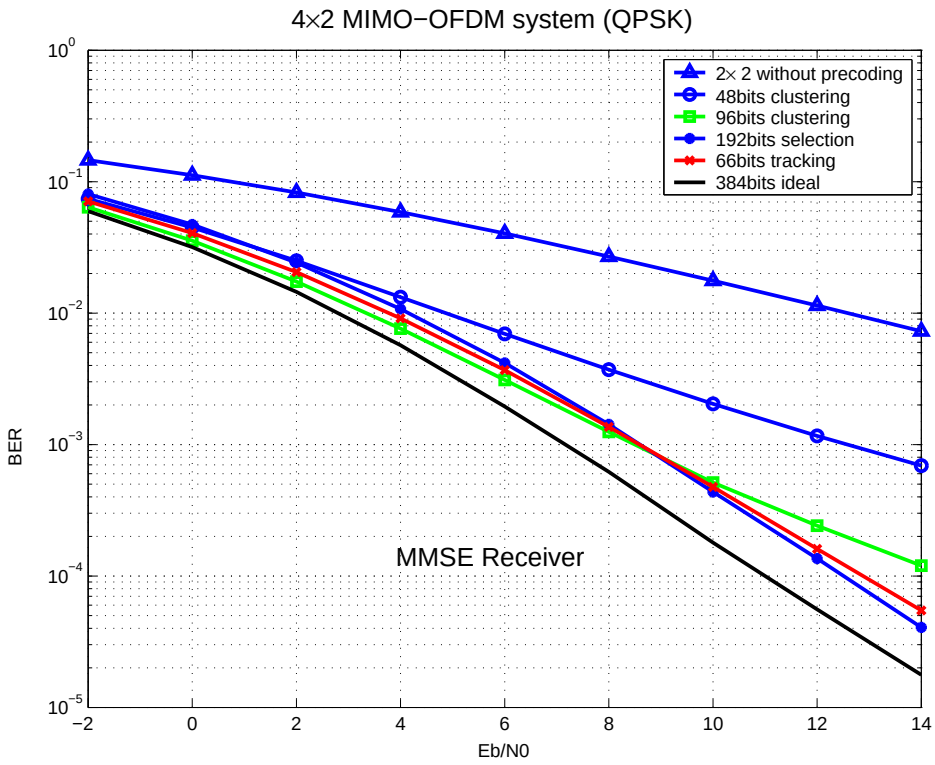


Figure 4.13: BER of a 4×2 MIMO-OFDM system with precoding.

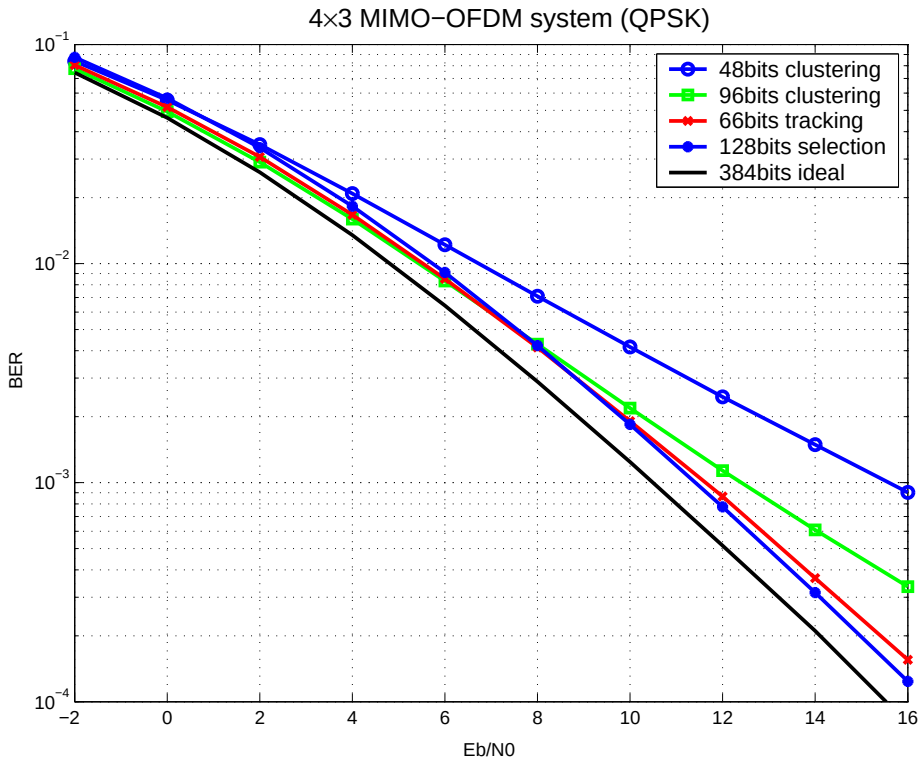


Figure 4.14: BER of a 4×3 MIMO-OFDM system with precoding.

CHAPTER V

CONCLUSIONS AND FUTURE RESEARCH WORK

5.1 Summary of Contributions

The work in the thesis focuses on the techniques to improve the reliability and capacity of OFDM-based wireless communication systems. The specific areas of our contributions are as follows:

1. ICI suppression for OFDM systems
2. Interference suppression for clustered OFDM
3. Clustered OFDM based anti-jamming modulation
4. Channel estimation for MIMO-OFDM
5. Precoding for MIMO-OFDM with limited feedback

In Chapter 2, We propose a frequency domain partial response coding for Suppressing ICI resulting from the Doppler frequency. We derive the ICI power for OFDM systems with PRC. Based on the general expression of the ICI power, the approximately optimal weights for PRC that minimize the ICI power are obtained. The numerical and simulation results show that PRC effectively reduces the error floor caused by Doppler frequency shift or carrier offset.

In Chapter 3, we address two techniques for clustered OFDM : interference suppression for clustered OFDM with adaptive antenna arrays and clustered OFDM based anti-jamming modulation.

For a polynomial-based parameter estimator for clustered OFDM with receive antenna arrays for interference suppression is proposed to combat severe leakage of the DFT based estimator. The developed estimator can be used in clustered OFDM systems with antenna arrays for interference suppression.

To calculate the weights of MMSE-DC in the clustered OFDM system with receive antenna arrays for interference suppression, instantaneous correlations of the received signals and channel responses corresponding to the desired signals have to be estimated. We propose a polynomial-based estimator to obtain the required parameters. Two issues are important for the polynomial-based estimator: the polynomial order and window size of the estimation. We study the impact of the polynomial order and window size on the estimation error. We develop an adaptive algorithm to obtain the optimal window size. With the adaptive algorithm, the polynomial-based estimator has no leakage and is robust to the channel statistics. Simulation results show that the developed algorithm improves performance of the clustered OFDM system with adaptive antenna array significantly.

For clustered OFDM used in military communications , we propose a clustered OFDM based spread spectrum system for anti-jamming. The proposed system has both advantages of clustered OFDM and spread spectrum. We analyze the performance for broadband and partial band jamming. For broadband jamming, we use the exponential correlation matrix to approximate the exact channel correlation matrix and obtain an approximate expression of the BER. The simulation results show that the approximate expression is very accurate for all SJR region. We also give the diversity and coding gain of the system that determine the error performance at high SJR. From the analytical results, high performance gain can be obtained for broadband jamming. However, the performance is degraded severely by the worst case partial band jamming. Our analysis shows that no diversity gain can be obtained at

the worst case jamming. To recover the performance loss, channel coding and interleaving are used. We evaluate the performance of the coded system using cutoff rate and upper bound. Two decoding schemes are investigated, hard-decision decoding and soft-decision decoding. Hard-decision decoding has low complexity and can recover most of the performance loss. To further improve the performance, soft-decision decoding is used. The analysis and simulation show that the soft-decision decoding can almost recover the performance loss. Since soft-decision decoding requires the jamming state information, we develop a simplified jamming state estimator. The simulation results show that the receiver with the estimated JSI performs very well over the receiver without JSI.

In Chapter 4, we investigate the OFDM system employing multiple antennas at both transmitter and receiver. The channel estimation and linear precoding for MIMO-OFDM systems are studied.

For MIMO-OFDM systems in correlated fading channels, we develop an MMSE channel estimator that makes full use of the spatial and frequency correlations. We also investigate the training sequence design for the estimator. By the conditions of the optimum training sequences, the training sequence design is only dependent on the transmit correlations. Then, we design the training sequences for two different cases. In the general cases, two optimum training sequences are proposed. For the special cases, the optimum training schemes are to assign the power over the eigenmodes of the dominant correlation matrix. The exact solution can be numerically calculated using Langrange multipliers. For high and low SNR region, we derive the close-form solutions.

For MIMO-OFDM systems in slow varying fading channels, partial channel information can be obtained through limited feedback. We propose to use linear precoding to achieve the diversity gain and array gain. The receiver selects the best precoding matrix from a codebook and conveys the index of the matrix to the transmitter. We

first investigate the codebook design and develop a random search algorithm. Then, we discuss the precoding matrix selection. Some commonly use selection criteria are introduced and the suboptimal selection criteria are proposed to reduce the computational complexity. In general MIMO-OFDM systems, the subchannels are highly correlated. Exploiting the property, we propose a scheme to reduce the required feedback bits. Previous work based on interpolation is only valid for single data stream transmission. We propose a subspace tracking based approach. The proposed approach is robust for multiple data stream transmission. Our approach do not need to recalculate the precoding matrices and can significantly reduce the feedback bits and selection complexity.

5.2 Future Research Work

MIMO-OFDM technique is a promising solution to high-data-rate wireless communications. In the thesis, techniques for channel estimation and linear precoding are proposed to increase the system performance. There are still challenging issues on MIMO-OFDM. Specifically, we will investigate the following issues:

- Theoretical analysis is needed for a better understanding of the precoded systems. With the theoretical analysis, we are able to determine some system parameters such as codebook size, step size so that the best performance/feedback bandwidth tradeoff can be achieved. In practice, wireless channel is time-varying and there is also feedback delay. Therefore, the precoding matrix selection algorithm should also consider these issues. The feedback rate is also determined by the characteristics of time-varying channel. Those issues are important for the realistic system design.

- Linear precoding for interference suppression: When multiple users are sharing the same channel, the signals from different users interfere each other. Like adaptive antenna array, multiple transmit antennas can also be used to suppress

interference. If the interference statistics can be tracked, then the receiver can choose the best precoding matrix for the transmitter to maximize the *signal to interference-plus-noise ratio*(SINR).

REFERENCES

[1] J. G. Proakis, *Digital communications,* 4rd ed., New York, McGraw-Hill, 2000.

[2] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Trans. on Comm.*, vol. COM-19, pp. 628-634, Oct. 1971.

[3] L. J. Cimini, Jr., "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. on Comm.*, vol. 33, pp. 665-765, July 1985.

[4] R. W. Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission", *Bell Syst. Tech. J.*, vol. 45, pp. 1775-1796, Dec. 1966.

[5] T. Cover and J. Thomas, *Elements of Information Theory,* New York, Wiley, 1991.

[6] G. L. Stüber, *Principles of Mobile Communication,* 2nd ed., Kluwer Academic, 2001.

[7] A. Paulraj, R. Nabar and D. Gore, *Introduction to Space-Time Wireless Communications,* Cambridge Univ. Press, May 2003.

[8] W. C. Jakes, *Microwave Mobile Communication,* 2nd ed., Piscataway, NJ, IEEE Press, 1994.

[9] S. Bhashyam, A. Sabharwal and B. Aazhang, "Feedback gain in multiple antenna systems," *IEEE Trans. on Comm.*, vol. 50, pp. 785-798, May 2002.

[10] S.A. Jafar, Sriram Vishwanath and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," *Proc. IEEE ICC 2001*, June 2001, pp. 2266 - 2270.

[11] N. Seshadri and J. H. Winters, "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity," *Proc. IEEE VTC*, 1993 pp. 508-511.

[12] J. H. Winters, "The diversity gain of transmit diversity in wireless systems with Rayleigh fading," *IEEE Trans. Vehicular Technology*, vol. 47, pp. 119-123, Feb. 1998.

[13] H. Jafarkhani, "A quasi-orthogonal space-time block code," *IEEE Trans. on Comm.*, vol. 49, pp. 1-4, Jan. 2001.

[14] A. F. Molisch and M. Z. Win, "MIMO systems with antenna selection," *IEEE Microwave Magazine*, vol. 5, pp. 46-56, Mar 2004.

[15] G. Foschini, G. Golden, R. Valenzuela, and P. Wolniansky, "Simplified processing for high spectal effciency wireless communication employing multi-element arrays," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1841-1852, Nov. 1999.

[16] M. Russell and G. L. Stüber, "Interchannel interference analysis of OFDM in a mobile environment," in *Proc. VTC'95,* 1995, pp. 820-824.

[17] P. Robertson and S. Kaiser, "The effects of Doppler spreads in OFDM(A) mobile radio systems," in *Prof. VTC'99-Fall*, 1999, pp. 329-333.

[18] P. Robertson and S. Kaiser, "Analysis of the loss of orthogonality through Doppler spread in OFDM systems," in *GLOBECOM'99*, vol. 1b, 1999, pp. 701-706.

[19] Y. (G.) Li and L. J. Cimini, Jr., "Bounds on the interchannel interference of OFDM in time-varying impairments communications," *IEEE Trans. on Comm.,* vol. 49, pp. 401-404, March 2001.

[20] B. Daneshrad, L. J. Cimini, Jr. and M. Carloni, "Clustered-OFDM transmitter implementation," *Seventh IEEE International Symposium on PIMRC'96.,* vol. 3, 1996, pp. 1064-1068.

[21] L. J. Cimini, Jr., B. Daneshrad and N. R. Sollenberger, "Clustered OFDM with transmitter diversity and coding," *IEEE, GLOBECOM '96.,* vol. 1, 1996, pp. 703-707.

[22] Y. (G.) Li and N. R. Sollenberger, "Clustered OFDM with channel estimation for high rate wireless data," *IEEE Trans. on Commun.*, vol. 49, pp. 2071-2076, Dec. 2001.

[23] J. Ahn and H. S. Lee, "Frequency domain equalization of OFDM signal over frequency nonselective Rayleigh fading channels," in *Electron. Lett.,* vol. 29, pp. 1476-1477, Aug. 1993.

[24] W. G. Jeon, K. H. Chang, and Y. S. Cho, "An equalization technique for orthogonal frequency division multiplexing systems in time-variant multipath channels," *IEEE Trans.on Comm.,* vol. 47, pp. 27-32, Jan. 1999.

[25] C. Muschallik, "Improving an OFDM reception using an adaptive Nyquist windowing," *IEEE Trans. Consumer Electron.,* vol. 42, pp. 259-269, Aug. 1996.

[26] J. Armstrong, "Analysis of new and existing methods of reducing intercarrier interference due to carrier frequency offset in OFDM," *IEEE Trans. on Comm.,* vol. 47, pp. 365-369, March 1999.

[27] S. Diggavi, N. Al-Dhahir, and A. Stamoulis, "Intercarrier interference in MIMO OFDM," In *Proc. IEEE ICC'2002*, Oct. 2002, pp.485-489.

[28] Y. Zhao and S. G. Häggman, "Intercarrier interference self-cancellation scheme for OFDM mobile communication systems," *IEEE Trans. on Comm.,* vol. 49, pp. 1185-1191, July 2001.

[29] Y. Zhao and S. G. Häggman, "Intercarrier interference compression in OFDM communication systems by using correlative coding," *IEEE Comm. Lett.*, vol. 2 pp. 214 -216, Aug.1998.

[30] H. Zhang and Y. Li, "Optimum frequency-domain partial response encoding in OFDM system," *IEEE Trans. on Comm.,* vol. 51, pp. 1064-1068, July 2003 / In *Proc. IEEE ICC'2003* May 2003, pp. 2025-2029.

[31] P. H. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. on Comm.*, vol. 42, pp. 2908-2914, Oct. 1994.

[32] S. Alam, "A general coding method to minimize inter-carrier interference in OFDM mobile communication systems," in *Proc. Int. Wireless and Telecomm. Symp./Exhibition (IWTS'97)*, vol. 1, Malaysia, May 1997, pp. 231-235.

[33] G. D. Forney, Jr., "Maximum Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference," *IEEE Trans. Information Theory*, IT-18, pp. 363-378, May 1972.

[34] G. L. Stüber, J. R. Barry, S. W. Mclaughlin, Y. G. Li, M. A. Ingram, T. G. Pratt, "Broadband MIMO-OFDM wireless communications," *Proceedings of the IEEE*, vol. 92, pp. 271-294, Feb. 2004.

[35] M. G. Robert, *Toeplitz and Circulant Matrices: A review*, http://ee.stanford.edu/ gray/toeplitz.pdf.

[36] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed., New York, McGraw-Hill, 1991.

[37] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ, Prentice-Hall, 1989.

[38] Y. (G.) Li, L. J. Cimini, Jr. and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading Channels," *IEEE Trans. on Comm.*, vol. 46, pp. 902-915, July 1998.

[39] J. Chuang and N. R. Sollenberger, "Beyond 3G: wideband wireless data access based on OFDM and dynamic packet assignment," *IEEE Commun. Magazine*, pp. 78-87, July 2000.

[40] G. Song and Y. Li, "Utility-based joint physical-mac layer optimization in OFDM," *IEEE, GLOBECOM'02.*, vol. 1, Nov. 2002, pp. 671-675.

[41] J. H. Winters, "Optimum combining in digital mobile radio with cochannel interference," *IEEE J. Select. Areas Commun.*, vol. sac-2, pp. 528-539, July 1984.

[42] J. H. Winters, J. Salz and R.D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. on Commun.*, vol. 42, pp. 1740-1751, Feb.-Apr. 1994.

[43] J. H. Winters, "Signal acquisition and tracking with adaptive arrays in the digital mobile radio system IS-54 with flat fading," *IEEE Trans. Veh. Technol.,* vol. 42, pp. 377-384, Nov. 1993.

[44] R. L. Cupo, G. D. Golden, C. C. Martin, K. L. Sherman, N. R. Sollenberger, J. H. Winters and P. W. Wolniansky,"A four-element adaptive antenna array for IS-136 PCS base stations," *IEEE 47th Vehicular Technology Conference,* vol. 3, 1997, pp. 1577-1581.

[45] Y. (G.) Li and N. R. Sollenberger, "Adaptive antenna arrays for OFDM systems with cochannel interference," *IEEE Trans. on Commun.*, vol. 47, pp. 217-229, Feb. 1999.

[46] H. Zhang and Y. (G.) Li, "Clustered OFDM with adaptive antenna arrays for interference suppression," In *Proc. IEEE ICC'2003* May 2003, pp. 2066-2070/*IEEE Trans. on Wireless Commun.*, to appear.

[47] I. E. Telatar, "Capacity of multiantenna Gaussian channels," *Eur. Trans. Commun.*, vol. 10, no. 6, pp. 585-595, 1999.

[48] G. J. Foschini and M.J. Gans, "On the limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311-335, March 1998.

[49] G. Raleigh and J. M. Cioffi, "Spatial-temporal coding for wireless communications," *IEEE Trans. Commun.*, vol. 46, pp. 357-366, 1998.

[50] S. Alamouti, "Space block coding: A simple transmitter diversity technique for wireless communications," *IEEE J. Select. Areas. Commun.*, vol. 16, pp. 1451-1458, Oct. 1998.

[51] V. Tarokh, H. Jafarkhani and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456-1467, July 1999.

[52] V. Tarokh, N. Seshadri and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744-765, Mar. 1998.

[53] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multielement antennas," *Bell Labs Tech. J.*, pp. 41-59, Autumn 1996.

[54] L. Zheng and D. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1073-1096, May 2003.

[55] Y. (G.) Li, N. Seshadri and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 461-471, March 1999.

[56] J. P. Kermoal, L. Schumacher, P. E. Mogensen and K. I. Pedersen, "Experimental investigation of correlation properties of MIMO radio channels for indoor picocell scenario," in *Proc. IEEE Veh. Technol. Conf.*, Boston, vol. 1, pp. 14-21, Sept. 2000.

[57] P. Kyritsi and D. C. Cox, "Correlation properties of MIMO radio channels for indoor scenarios," in *proc. Signal Systems and Computers 35th Asilomar Conf.*, vol. 2, 2001.

[58] M. K. Simon, J. K. Omura, R. A. Scholtz and B. K. Levitt, *Spread Spectrum communications*, Vols I&II, Rockville, MD: Computer Sience Press, 1985.

[59] K. Cheun, K. Choi, H. Lim and K. Lee, "Antijamming performance of a multi-carrier direct-sequence spread-spectrum system," *IEEE Trans. on Comm.*, vol. 33, pp. 1781-1784, Dec 1999.

[60] S. Zhou, G. B. Giannakis and A. Swami, "Digital multi-carrier spread spectrum versus direct sequence spread spectrum for resistance to jamming and multi-path," *IEEE Trans. on Comm.*, vol. 50, pp. 643-655, Apr. 2002.

[61] E. Lance and G. K. Kaleh, "A diversity scheme for a phase-coherent frequency-hopping spread-spectrum system," *IEEE Trans. on Comm.*, vol. 45, pp. 1123-1129, Sep. 1997.

[62] H. Zhang and Y. (G.) Li, "Anti-jamming properties of clustered OFDM spread spectrum systems for dispersive channels," In *Proc. Milcom'2003* Oct. 2003/*IEEE Trans. on Commun.*, submitted.

[63] J.-J. van de Beek, O. S. Edfors, M Sandell, S. K. Wilson and P. O. Borjesson, "On channel estimation in OFDM systems," *Proc. 45th IEEE Vehicular Technology Conf., Chicago, IL*, July 1995, pp. 815-819.

[64] O. Edfors, M. Sandell, J.-J. van de Beek, S. K. Wilson and P. O. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. on Commun.*, vol. 46, pp. 931-939, July 1998.

[65] J. Fan and I. Gijbels, "Data-driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation," *J. Royal Statist. Soc.* vol. 57, pp. 371-394, 1995.

[66] J. Fan and I. Gijbels, *Local Polynomial Modelling and its Applications*, Chapman and Hall, London, 1996.

[67] X. Wang and K. J. R. Liu, "Channel estimation for multicarrier system using a time-frequency polynomial model," *IEEE Trans. on Commun.*, vol. 50, pp. 1045-1048, July 2002.

[68] D. K. Borah and B. T. Hart, "Frequency-selective fading channel estimation with a polynomial time-varying channel model," *IEEE Trans. on Commun.*, vol. 47, pp. 360-364, Mar. 1999.

[69] D. K. Borah and B. T. Hart, "A robust receiver structure for time-varying, frequency-flat rayleigh fading channels," *IEEE Trans. on Commun.*, vol. 47, pp. 862-873, June 1999.

[70] E. W. Cheney, *Introduction to the Approximation Theory*, New York: McGraw-Hill, 1966.

[71] K. S. Miller, *Complex Stochastic Processes: An Introduction to Theory and Application*. Reading, MA: Addison-Wesley, 1974.

[72] R. K. Mallik, "On multivariate Rayleigh and exponential distributions," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1499-1515, June 2003. Page(s): 1499- 1515

[73] Z. Wang and G. B. Giannakis, "A simple and general parameterization quantifying performance in fading channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 1389-1398, Aug. 2003.

[74] L. E. Blumenson and K. S. Miller, "Properties of generalized Rayleigh distributions," *Ann. Math. Statist.*, vol. 34, pp. 903-910, 1963.

[75] G. K. Karagiannidis, D. A. Zogas and S. A. Kotsopoulos, "On the multivariate nakagami-m distribution with exponential correlation," *IEEE Trans. on Comm.*, vol. 51, pp. 1240-1244, Aug. 2003.

[76] G. K. Karagiannidis, D. A. Zogas and S. A. Kotsopoulos, "Performance analysis of triple selection diversity over exponentially correlated nakagami-m fading channels," *IEEE Trans. on Comm.*, vol. 51, pp. 1245-1248, Aug. 2003.

[77] S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 5th ed. New York: Academic, 1994.

[78] J. P. Odenwalder, "Optimum decoding of convolutional codes," Doctoral dissertation, Univ. Calif., Los Angeles, 1970.

[79] H. Bölcskei, M. Borgmann and A. J. Paulraj, "Impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM," *IEEE J. Select. Areas. Commun.*, Vol. 21, pp. 427-439, Apr. 2003.

[80] J. Tan and G. L. Stüber, "Multicarrier delay diversity modulation," *IEEE GLOBECOM'03*, vol. 3, pp. 1633-1637, Dec. 2003.

[81] H. Bölcskei, M. Borgmann and A. J. Paulraj,"Space-frequency coded MIMO-OFDM with variable multiplexing-diversity tradeoff," *Proc. IEEE Int. Conf. Communications*, Anchorage, AK, pp. 2837-2841, May 2003.

[82] Y. (G.) Li "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. Wireless Communications*, vol. 1, pp. 67-75, Jan. 2002.

[83] I. Barhumi, G. Leus and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Processing*, vol. 51, pp. 1615-1624, June 2003.

[84] L. Schumacher, K. I. Pedersen, and P. E. Mogensen, "From antenna spacings to theoretical capacities-guidelines for simulating MIMO systems," in *Proc. PIMRC Conf.*, vol. 2, pp. 587-592, Sept. 2002.

[85] H. Bölcskei, D. Gesbert and A. J. Paulraj, "On the capacity of OFDM based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, pp. 225-234, Feb. 2002.

[86] J. H. Kotecha and A. M. Sayeed, "Transmit Signal Design for Optimal Estimation of Correlated MIMO Channels Kotecha," *IEEE Trans. Signal Processing*, vol. 52, pp. 546-557, Feb. 2004.

[87] G. Barriac and U. Madhow, "Space time communication for OFDM with implicit channel feedback," submitted for publication.

[88] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Upper Saddle River, NJ: Prentice-Hall PTR, 1998.

[89] H. Shin and J. H. Lee "Capacity of multiple-antenna fading channels: spatial fading correlation, double scattering, and keyhole." *IEEE Trans. Inform. Theory*, vol. 49, pp. 2636-2647, Oct. 2003.

[90] C. Chuah, D. Tse, J. Kahn, and R. Valenzuela "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, vol. 48, pp. 637-650, March 2002.

[91] M. Godavarti, T. L. Marzetta, and S. Shamai, "Capacity of a mobile multiple-antenna wireless link with isotropically random Rician fading," *IEEE Trans. Inform. Theory*, vol. 49, pp. 3330-3334, Dec. 2003.

[92] R. U. Nabar, H. Bölcskei, and A. J. Paulraj, "Diversity performance of Ricean MIMO channels," *ITG Workshop on Smart Antennas*, Munich, Germany, Mar. 2004, to appear.

[93] V. Erceg, L. Schumacher, P. Kyritsi, A. Molisch, D. S. Baum, A. Y. Gorokhov, C. Oestges, Q. Li, K. Yu, N. Tal, B. Dijkstra, A. Jagannatham, C. Lanzl, V. J. Rhodes, J. Medbo, D. Michelson, M. Webster, E. Jacobsen, D. Cheung, C. Prettie, M. Ho, S. Howard, B. Bjerke, L. Jengx, H. Sampath, S. Catreux, S. Valle, A. Poloni, A. Forenza, R. W. Heath, *IEEE P802.11 Wireless LANs: TGn Channel Models* (IEEE 802.11-03/940r1), 44 p., 2004.

[94] L. Schumacher, "WLAN MIMO channel Matlab program," download information: http://www.info.fundp.ac.be/∼lsc/Research/IEEE_802.11_HTSG_CMSC/distribution_terms.html

[95] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1423-1436, Oct. 1998.

[96] D. J. Love, R. W. Heath, Jr., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2735-2747, Oct. 2003.

[97] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple-antenna systems," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2562-2579, Oct. 2003.

[98] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, R. Urbanke, "Systematic design of unitary space-time constellation," *IEEE Trans. Inform. Theory,* vol. 46, pp. 1962-1973, Sep. 2000.

[99] D. J. Love and R. W. Heath, Jr., "Limited feedback precoding for spatial multiplexing systems using linear receivers," in *Proc. IEEE Mil. Comm. Conf.*, Oct. 2003.

[100] D. Love and R. W. Heath Jr., "Limited feedback precoding for spatial multiplexing systems," *IEEE GLOBECOM'03* vol. 4, Dec. 2003, pp. 1857-1861.

[101] J. Choi and R. W. Heath, Jr., "Interpolation based transmit beamforming for MIMO-OFDM with limited feedback," *Proc. IEEE ICC*, vol. 1, June 2004, pp. 249-253.

[102] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian spaces," *Experimental Mathematics*, vol. 5, pp. 139-159, 1996.

[103] D. Agrawal, T. J. Richardson, and R. L. Urbanke, "Multiple-antenna signal constellations for fading channels," *IEEE Trans. Inform. Theory.*, vol. 47, pp. 2618-2626, Sept. 2001.

[104] W. M. Boothby, *An Introduction to Differential Manifolds and Riemannian Geometry*, 2nd ed. San Diego, CA: Academic, 1986.

VITA

Hua Zhang was born in January 1975 in Haimen, Jiangsu Province, P. R. China. He received his B.S. and M.S. degrees from the Department of Radio Engineering, Southeast University, Nanjing, China in 1998 and 2001, respectively. From September 1998 to March 2001, he was with the National Mobile Communications Research Laboratory at Southeast University as Graduate Research Assistant. From August 2001 to December 2004, he was a Graduate Research Assistant at the Information Transmission and Processing Laboratory at the Georgia Institute of Technology. He completed his Ph.D. in Electrical and Computer Engineering from Georgia Institute of Technology in December 2004. His research interests are signal precessing for wireless communications, including OFDM, interference suppression, space-time processing.